DOCUMENT RESUME

ED 059 795 PS 005 502

AUTHOR Taebel, Donald K.
TITLE The Effect of Various Instructional Modes on Children's Performance of Music Concept Tasks.
INSTITUTION Concordia Teachers Coll., Seward, Nebr.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-0-F-072
PUB DATE Dec 71
GRANT OEG-6-70-0038(509)
NOTE 112p.

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS Age Differences; *Children; *Concept Formation; Effective Teaching; Elementary School Students; Individual Differences; Kindergarten Children; Measurement Instruments; Mental Development; *Music Education; Objectives; Private Schools; Public Schools; Research; Stimulus Devices; Student Evaluation; *Task Performance; *Teaching Techniques; Tests

ABSTRACT

The development of certain concepts of music in kindergarten, first, and second grade children was researched. Objectives of the study were: (1) To measure the effect of four instructional modes on the performance of tasks which embodied certain musical concepts; (2) To measure the performance differences between age levels; and (3) To measure and evaluate young children's concepts of louder, faster, higher, and shorter as elicited by the measuring instrument. Two forms of a measuring instrument were developed. Each form contained several tests which measured a specific concept. Each test consisted of 20 items or tasks with each task displaying positive and negative instances of the concept. The subject was asked to identify the "different," i.e. negative, instance of each set. Two independent samples were drawn from public and parochial schools of Lincoln, Nebraska, and the surrounding region. Instructional modes used were designated discovery, verbal cue, verbal response, and motor response. There were significant differences in performance between tests which embodied different concepts. It was concluded that the level of performance is a function of the stimulus features of the task and the accessibility of the concept. Other conclusions are: (1) Performance on tasks is partly a function of age; and (2) There is usually no significant effect of the instructional mode when a concept is not readily available as in the case of higher. Further research is recommended. (Author/CK)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

OE-BR
PS

ED 059795

Final Report

Project No. O-F-072

Grant No. OEG-6-70-0038 (509)

The Effect of Various Instructional Modes on Children's Performance of Music Concept Tasks

Donald K. Taebel

Concordia Teachers College

Seward, Nebraska

December, 1971

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

PS 005502

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

CONTENTS

Title Page i
Table of Contents ii
List of Illustrations iii
List of Tables iv
Acknowledgments vi

Summary of the Study 1

The Problem and Objectives of the Study 3

Some Theoretical Aspects of Concept Development 8
- The Nature of Concept Development 8
- Interrelationship between Perceptual and Conceptual Development 17
- Relationship between Language and Conceptual Development 25

Development and Implementation of Research 32
- Development of the Long Form 32
- Development of the Short Form 50
- The Sample 52
- Research Design 53

Results of the Study 55
- Hypothesis One: Task Variables 55
- Hypothesis Two: Age Variables 59
- Hypothesis Three: Instructional Variables 66
- Reliability 74
- Validity 75
- Interpretation of the Results 77

Summary and Conclusions 83
- Implications for Music Education 86
- Recommendations for Further Research 87

References 89

Appendixes 96

ILLUSTRATIONS

Figure 1. Concept Task Where Pitch Is Relevant 33

Figure 2. Visual Model of Concept Task 35

Figure 3. Tempo and Volume Changes Across Melodic Patterns 36

Figure 4. Sets of Triangles Illustrating Varying Degrees of Discriminability 40

Figure 5. Differing Estimates of Volume Changes When There Was No Change 41

Figure 6. Differing Estimates of Volume Changes When There Was a Change of Two Decibels 41

Figure 7. Differing Estimates of Volume Changes When There Was a Change of Four Decibels 42

Figure 8. Differing Estimates of Volume Changes When There Was a Change of Six Decibels 42

Figure 9. Differing Estimates of a Change in Tempo When There Was No Change 43

Figure 10. Differing Estimates of a Change in Tempo When There Was a Change of Six MM Units 43

Figure 11. Differing Estimates of a Change in Tempo When There Was a Change of Twelve MM Units 44

Figure 12. Differing Estimates of a Change in Tempo When There Was a Change of Eighteen MM Units 44

Figure 13. Volume and Tempo Changes in Sample Task 45

Figure 14. Tempo and Volume Changes in Sample Task 45

Figure 15. Summary of Percentage of Correct Responses On Each of the Sequences by Age Group 59

Figure 16. Ogive on Volume Tasks 61

Figure 17. Ogive on Tempo Tasks 62

Figure 18. Ogive on Pitch Tasks 64

Figure 19. Ogive on Duration Tasks 65

TABLES

Table 1. Tasks Involving Conceptual Behavior 16

Table 2. Differing Estimates of Pitch Changes 40

Table 3. Percentage Who Identified Changes of Pitch 47

Table 4. Division of Sample by Grade and Form 53

Table 5. Division of Sample by Grade, Mode, and Form 54

Table 6. Probability of Observed Means on Volume Tasks 56

Table 7. Probability of Observed Means on Tempo Tasks 57

Table 8. Probability of Observed Means on Pitch Tasks 57

Table 9. Probability of Observed Means on Duration Tasks 58

Table 10. Means and S.D. on Volume Tasks 60

Table 11. Analysis of Variance on Volume Tasks 60

Table 12. Means and S.D. on Tempo Tasks 61

Table 13. Analysis of Variance on Tempo Tasks 62

Table 14. Means and S.D. on Pitch Tasks 63

Table 15. Analysis of Variance on Pitch Tasks 63

Table 16. Means and S.D. on Duration Tasks 64

Table 17. Analysis of Variance on Duration Tasks 65

Table 18. Means for Each Mode - Volume Tasks 67

Table 19. Analysis of Variance of Column Means - Volume 67

Table 20. Verbal Reports on Volume Tasks 68

Table 21. Means for Each Mode - Tempo Tasks 69

Table 22. Analysis of Variance of Column Means - Tempo 69

Table 23. Verbal Reports on Tempo Tasks 70

Table 24. Means for Each Mode - Pitch Tasks 70

Table 25. Analysis of Variance of Column Means - Pitch 71

Table 26. Verbal Reports on Pitch Tasks 72

Table 27. Means for Each Mode - Duration Tasks 72

Table 28. Analysis of Variance of Column Means - Duration 73

Table 29. Verbal Reports on Duration Tasks 73

Table 30. Reliability of the Tests 75

Table 31. Correlation of Teacher Ratings and Test Results 76

Table 32. Test Correlations 77

ACKNOWLEDGMENTS

Many people have contributed to this project, unfortunately, most of them will have to remain anonymous. Thanks must be given to the children, teachers, and principals of the following schools: Cowan, Westchester Lutheran, and Covenant Presbyterian schools of Los Angeles, California; Pyrtle, Saratoga, Prescott, and Calvary elementary schools; and Ruth Staples Laboratory Nursery School of Lincoln, Nebraska; Seward, Staplehurst, Goehner, Waco, and Beaver Crossing elementary schools; and St. John Lutheran School of Seward, Nebraska. The greatest debt is owed to the children who so willingly participated in the interview sessions. The principals and teachers must also be commended for their helpfulness and sympathetic understanding.

Dr. Raymond Haggh of the University of Nebraska kindly gave his time and expertise in producing the melodic patterns on the Moog Synthesizer. Mr. Ken Husky of the University of Nebraska Audio-Visual Center was most generous in providing technical advice and in processing of the tapes. Dr. Vance Hinrichs and Dr. J. D. Weinhold of Concordia Teachers College must be cited for their valuable assistance with the statistical analysis. A major contributor to the project was Professor Don Sylwester who was involved with the development of the instrument in terms of quantifying the stimulus patterns, and who wrote the computer programs and provided keen insight into the problems of statistical analysis.

In the background, but nonetheless of vital importance, at each stage of the project were the love, concern, and encouragement of my wife and children, Alleane, Jeff, and Kim.

Summary of the Study

Purpose and Objectives

The purpose of the research was to study the development of certain concepts of music in kindergarten, first, and second grade children. The objectives of the study were: (1) to measure the effect of four instructional modes on the performance of tasks which embodied certain musical concepts, (2) to measure the performance differences between age levels, and (3) to measure and evaluate young children's concepts of louder, faster, higher, and shorter as elicited by the measuring instrument.

Method

Since no instrument was available which suited the objectives of the study, two forms of the instrument were developed. Each form contained several tests which measured a specific concept. Each test consisted of 20 items or tasks with each task displaying positive and negative instances of the concept. For example, in a Long form task three instances may be at the same tempo while the fourth may be faster. The subject was asked to identify the "different", i.e. negative, instance of each set.

Two independent samples were drawn from public and parochial schools of Lincoln, Nebraska and the surrounding region. The Long form was given to 126 children, the Short form to 260 children. The children were randomly selected from their classes and placed into four nearly-equal treatment groups. All groups received the same orientation instructions; however, subsequent instructions varied according to the treatment group. The instructional modes were designated discovery, verbal cue, verbal response, and motor response.

Results and Conclusions

The results of the experiment showed:

There were significant differences in performance between tests which embodied different concepts. It was concluded that the level of performance is a function of the stimulus features of the task and the accessibility of the concept. The children appeared to use a concept of volume very readily. Performance on tempo tasks was slightly lower. The concept of shorter did not appear to be consistently used. The concept of higher was poorly used and frequently confused with louder.

There were significant differences in performance on all tests between the age groups, especially between kindergarten and first grade levels. It was concluded that performance on these tasks is partly a function of age. The discontinuity between the kindergarten and first grade performance was interpreted as a possible manifestation of the shift from pre-operational to operational thought which, according to Piaget's theory, occurs about age seven.

The effect of the instructional mode on performance was not significant in a consistent manner but varied with age and the kind of conceptual task. It was concluded that there is usually no significant effect of the instructional mode when a concept is not readily available as in the case of higher. Neither is there a significant effect when the concept is well developed as in the case of louder. On the other hand, a verbal mode of instruction, either a verbal cue or verbal response which is reinforced, is effective when concept development is at some intermediate stage as in the case of faster.

It was recommended that further research be done on the relation between verbal modes of instruction and non-verbal modes at various age levels, particularly the pre-school level. Teachers should be particularly careful in their choice of words in describing musical phenomenon. This is especially important with the terms" higher" and "louder" since children frequently confuse the two.

CHAPTER ONE: THE PROBLEM AND OBJECTIVES OF THE STUDY

Introduction to the Problem

The development of concepts in the human being is held by many psychologists and educators to be of paramount importance in education (Gange, 1965). Woodruff (1970, p. 51) has forthrightly declared, "Conceptual development is not simply one of several possible things of interest to an educator; it is rather the essence of concerns of an educator." Over the past several years music educators have shown a growing interest in the development of concepts in children. This was noted by Carlsen (1969, p. 8) who remarked, "Another area for musical learning research is concept formation or conceptual behavior. This is particularly important now since, historically, we have been principally concerned with the development of skills." Further evidence of this concern for the development of concepts is found in a number of recent school music series (Watters, Wersen, Hartshorn, McMillan, Galup, Beckman, 1965; Sur, Tolbert, Fischer, McCall, 1967; Leonhard, Krone, Wolfe, Fullerton, 1967). This same emphasis on concept development is found in music education texts such as The Study of Music in the Elementary School - A Conceptual Approach (Gary, 1967) or those of Aronoff (1969), Cheyette and Cheyette (1969), Marsh (1970), and Nye and Nye (1970).

Although there is a general recognition of the importance of concept development, information about the kinds of concepts which children have at a particular age and the processes by which they are acquired is yet largely a matter of intuition and impressionistic evaluation. According to Andrews and Deihl (1968, p. 2), "An apparent gap exists between opinions of leading music educators regarding concept-centered curriculum content and the identification of a body of knowledge on children's concepts of musical elements." Aronoff (1969, p. 29) has pointed out that the music teacher will have to make an educated guess concerning the mode of instruction she uses since "there are still many unknowns in the processes through which the child actually sees relationships, categorizes, discriminates, and generalizes." Shuter (1969) has suggested that knowledge of what tasks children at various stages of development can be expected to undertake would be valuable to music education. In sum, there is, at present, little experimentally founded data on (1) the music concepts which a child has at a particular age, and (2) the modes of instruction which are suitable for developing concepts.

Statement of the Problem

The problem of the study was to determine the effect of various instructional modes on children's performance of music concept tasks.

Objectives of the Study

The objectives of the study were as follows:

1. To determine the effect of four instructional modes on the performance of music concept tasks.
2. To examine the relationship between age and performance on concept tasks.
3. To analyze young children's concepts of louder, faster, higher, and shorter as elicited by tasks in which they sort musical patterns.

Background of the Study

Although teaching methodology in music has been mainly pragmatic rather than experimentally or theoretically based, there has been recently a serious effort to relate instructional strategies to certain theoretical principles of child development. Among the learning theories available, the cognitive-developmental theories of Piaget, Bruner, and Mursell are considered to hold the most promise for undergirding methods of music instruction (Carlsen, 1969; Aronoff, 1969; Nye, 1970; Larson, Bode, 1971).

Perhaps the most attractive aspect of cognitive-developmental theory for music educators is the acknowledgment of non-verbal behaviors as contributors to cognitive development. Intelligence, as understood by Piaget, should not be restricted to verbal or other symbolic behavior, but also includes motor and imaging behaviors, or, to use Bruner's terms, the enactive and iconic modes. The cognitive-developmental theories of Piaget and Bruner rest on the assumption that intellectual structures are developed through an interaction between the child and his environment. The chief forms which interaction takes in the early years of life are enactive and iconic (Kohlberg, 1969). "Piaget claims that object construction is dependent upon the coordination of polysensorial input-stimulation from visual, kinesthetic, and proprioceptive systems" (Charlesworth, 1969 p. 19).

An examination of the aforementioned music series and music education texts reveals an awareness of non-verbal activities as a means of developing concepts. A clear example of Piagetian theory being utilized to teach musical concepts is found in the work of Carabo-Cone (1969). According to Aronoff (1969, p. 8), "Music educators have intuitively emphasized the non-verbal aspects of a child's experience, but until now have perhaps failed to realize fully the musical learning which can take place in these ways."

If the above theoretical principles concerning cognitive development are accepted, it follows that a researcher seeking to diagnose the development of children's concepts must allow some non-verbal behavior as a valid indicator of concept attainment. Gange (1965) has explained that concept learning involves an internalization process by which experiences are represented. The crucial test of concept attainment is the generalized response to some common, abstract property. As far as the response is concerned, some form of non-verbal behavior is commonly acceptable because the child behaves as if he had a verbal explanation for his action. Aronoff (1969) stated that a concept is best treated as a hypothetical construct to explain observed facts. A person's concept cannot be directly observed; but it can sometimes be inferred from observation of his behavior.

Even though there are ample grounds for accepting non-verbal behavior as evidence of concept attainment, it would be a mistake to conclude that only non-verbal activities are appropriate for the development of concepts. The role of verbal instruction has been basic to pedagogy for so long that it cannot be dismissed in a cavalier manner. The relationship between concept learning and verbal instruction is both intimate and complex and has provoked much research (Bourne, 1966). Such is the case with the present experiment.

Related Research

Although Piaget's theory of intellectual development has spawned a host of studies on the development of children's concepts of natural science and mathematics, there has been little work done on the formation of music concepts. Aileen Kidd (1966) has suggested that a fruitful line of research may be developed using certain techniques of Piaget together with modern statistical procedures. Using Piagetian theory Marilyn Pflederer Zimmerman has conducted several studies on the development of music concepts. In her first study (Pflederer, 1964) melodic patterns were arbitrarily modified and five and eight year old children were asked to offer verbal explanations of what kind of change had been made. From these verbal remarks an estimate was made of the child's ability to conserve a common element across the tonal changes. A later study (Pflederer, 1968) used a similar research strategy but added more controls, increased the size of the sample, and broadened the age range. More recently Zimmerman (1970) has measured the effect of instructional modes on conceptual behavior. Additional work has been reported by Andrews and Deihl (1970) who have developed a technique for identifying music concepts of children in the intermediate grades using both verbal and non-verbal procedures.

PS 005502

The investigator has been engaged in research since 1969 on concept development and has studied some 90 children between the ages of five and nine years of age. During this time certain experimental procedures have been developed which have been incorporated into the present study. One of the intriguing aspects of this preparatory research was the relationship between verbal instruction and concept development. This has become the central purpose of this study.

Null Hypothesis

1. There is no significant difference between the observed mean on any task sequence and a mean which is attained only by chance.
2. There is no significant difference between the means attained by different age groups on any task sequence.
3. There is no significant difference between the means attained by groups which are given different modes of instruction.

Assumptions

1. The ability to classify tonal materials may be elicited by asking children to sort various melodic patterns into sets which "go together" and to exclude the "different" pattern.
2. Children who correctly categorize melodic patterns with consistency are demonstrating conceptual behavior.
3. Categorization is possible even though a subject cannot offer a verbal explanation for his choice.
4. A task which elicits concept behavior involves the presentation of materials which contain at least one common element but may vary in a number of features as well.

Organization of the Remainder of the Study

Chapter

II. Some Theoretical Aspects of Concept Development

The characteristics of concept development

The relationship between perception and concept development

The relationship between language and concept development

III. Development and Administration of Research Procedures
  - Development of the Long form instrument
  - Administrative procedures
  - Development of the Short form instrument
  - The sample
  - Research design

IV. Results of the Study
  - Results related to the research hypotheses
  - Reliability and validity
  - Interpretation of results

V. Summary, Conclusions, and Recommendations
  - Implications for music education
  - Recommendations for further research

References

Appendixes

CHAPTER II: SOME THEORETICAL ASPECTS OF CONCEPT DEVELOPMENT

The first purpose of this chapter was to provide a theoretical background of the experiment at hand. Such a background was considered valuable in providing a rationale for the research design and instrumentation. Furthermore, the issues to be considered were selected for their relevance to the research hypotheses and the assumptions as stated in chapter one. For the most part, this review of theoretical aspects was based on the work of cognitive-developmental psychologists.

Although some experimental research proceeds without explicit formulations of the theoretical perspective, the importance of some kind of theoretical orientation in the formulation and interpretation of an experiment should not be minimized. "It must be remembered that the ultimate business of research is not just to collect data but to find orderliness in nature. Experiments must be fused into meaningful generalities by logic, speculation, and theory" (White, 1963, p. 216). Theory is even more important in research connected with concept development for as Flavell (1970, p. 990) remarks, "Nowhere do facts without theory stand in lonelier isolation than in the area of concept development."

The second purpose of the chapter was to provide some basis for subsequent analysis and interpretation of the results.

The chapter was organized around three problems of conceptual development. First, the problem of defining and differentiating concept development was treated; secondly, the problem of the relationship between perception and conceptualization was outlined; and, thirdly, the problem of the interrelationship of conceptual development and verbal behavior was considered. In reality, these three types of behavior - perception, conceptualization, and verbalization - function together. Even though behavior is basically continuous, for the sake of efficient analysis, it is handy to have a finite set of behavioral categories at our disposal. It should be understood, however, that categories of behavior overlap and are not mutually exclusive (Bourne, 1966).

I. THE NATURE OF CONCEPT DEVELOPMENT

What is a Concept?

The search for a satisfactory definition of the term "concept" is a lexicographer's nightmare (Flavell, 1970). There are numerous definitions of the term each of which vary in their meaning or which emphasize a different aspect of behavior. For some the term is often used synonomously with "idea". Thus, Woodruff defines a concept as "a relatively complete and meaningful idea" (Gary, p. 2). Others (Watters, et. al., 1965, p. VI) have viewed the concept as some kind of image such as in the following definition: "Tonal concepts are mental images of tones that remain in the mind."

The difficulty with both of these definitions is that they do little more than substitute one term for the other. Unfortunately, the terms "idea" and "image" are themselves very difficult to define and resorting to a dictionary only proves the point since the definition of one of the words is in terms of the other. The use of the term "image" with respect to musical concepts is particularly troublesome since it evidently has been borrowed from the field of visual perception. Instead of laboring over a concise, consensually valid statement of what a concept is, it may be more instructive to describe what a concept does.

Functions of Concepts

Concepts are useful to the individual in that they simplify the experiential world. Concepts have the essential function of reducing the complexity of sensory input to manageable proportions (Flavell, 1970). Bruner, Goodnow, and Austin (1956) believe that concepts function to categorize events and objects thereby serving to cut down the diversity of objects and events that must be uniquely dealt with by an organism of limited capacities. Much, if not most, of the interaction between an individual and the environment involves dealing with classes or categories of things rather than particular events or objects (Bourne, 1966). For example, instead of responding in different ways to each of the twelve major scales, a person can respond in a general way to the class of scales which are major. This type of response is known as an equivalence response. Kendler (1961, p. 447) refers to such a response in his definition of a concept as a "common response to dissimilar stimuli." His argument is that a conceptual response is one in which a similar or identical reaction occurs to different environmental inputs. The equivalence response view of conceptual functioning indicates a semantic connection between concept and such terms as class, category, and set inasmuch as these terms also imply an act of drawing together disparate elements for response purposes (Flavell, 1970). Closely related to Kendler's definition is the following: "A concept extracts a common meaning from a diverse array of experiences" (CRM, 1971).

Another function of concepts is that their use saves time for the individual. "Having concepts saves the organism time in dealing with the world. Old and new inputs usually can be quickly identified and appropriately reacted to by virtue of being assimilable to existing concepts" (Flavell, 1970).

Concepts are useful in that they help an individual relate present input to the past since

> to know that a perceived object is an instance of a particular class is to know a great deal else that is not immediately perceptible; it is to know

> virtually all that it could do to you or you to it, because this information is implied in its class membership (Flavell, 1970, p. 986).

For example, to know that a piece of music is the first movement of a classical symphony activates a host of concepts related to the form of the first movements of classical symphonies. These concepts function in guiding and supplementing the immediate perceptual activity. Obviously, concepts do not function in isolation one from another. In general, about the most certain prediction one can make regarding the activation of any concept is that numerous others will be immediately activated as a consequence. The process of conceptual development proceeds through the development of interrelationships between concepts to form complex systems. As concepts become more complex, they will oftentimes be formed from other, more-simple, concepts. In a somewhat comparable manner a dictionary defines the unfamiliar, "harder" word in terms which are assumed to be familiar and more concrete. For example, the term "cadence" is defined by reference to such other concepts as chord, tonic, and dominant and their relationships to each other. The concept of a cadence, in turn, becomes a defining attribute of such other concepts as phrase, period, and more complex musical forms. A stimulus event which activates a particular concept sets off a chain reaction by means of which related concepts are brought into service in processing the stimulus input. Because of this network of concepts, the raw, sensory input is organized, modified, and transformed to a higher level than the immediate perception. The function of concepts as an information processing system is described by Hunt, Harvey, and Schroder (1961, p. 1):

> A concept is a system of ordering that serves as the mediating linkage between the input side (stimulus) and the output side (response). In operating as a system of ordering, a concept may be viewed as a categorical schema, an intervening medium, or a program through which impinging stimuli are coded, passed, or evaluated on their way to response evocation.

Concisely stated, a concept is a system for classifying stimuli (CRM, 1970). In sum, concepts function in three important ways:

1. They identify and classify input information.
2. They save time.
3. They make it possible for the individual to "go beyond the information given" in adapting to stimulus input.

Differences between Concepts

Although all concepts perform the same functions, they differ also in a number of significant ways. These differences between concepts fall into two distinguishable categories.

1. Concepts differ in their public, conventional qualities. It is possible to consider concepts as objective entities detached from an individual's personal experience. In so doing, concepts are viewed as end-products of developmental processes through which they come to have accepted public definition. The second way in which concepts differ is in the processes by which they are formed and used. It is one thing to know the public meaning of a concept but quite another to be aware of the individual, private processes which may be involved in conceptual behavior. Whereas, the first kind of difference is of interest to the philosopher and lexicographer, the second kind of difference is of considerable interest to the developmental psychologist and educator (Flavell, 1970).

Let us first consider the differences between concepts on the basis of their public definitions. The dimensions on which concepts vary are their attributes, rules, abstractness, generality, precision, and power.

Attribute Differences Concepts differ in the kind and number of attributes which serve to define them. Some object or event is an instance of a concept if, and only if, it possesses the set of attributes given in the definition of the concept. Such attributes are the defining or critical attributes; there are usually other, non-defining attributes as well in the stimulus situation. These kind of attributes are irrelevant to the definition. For example, the concept "clarinet" has certain defining attributes such as shape, mechanism, reed, timbre, and other irrelevant attributes such as color or material. The concept "trombone" is different in that it has other relevant attributes which define it.

Attributes can be almost anything that humans can perceive or think of although they generally are of two types.

(1) A perceptible or measurable physical property such as color, texture, or shape may be considered a structural attribute. Each musical instrument has particular, distinctive structural attributes.

(2) Attributes which are less concrete and fixed and refer to an action or use are transformational attributes. These attributes describe what the object could do or become under certain circumstances. For example, an orange besides having structural attributes, also has the transformational attribute of being edible. In like manner, the timpani has structural attributes but also has the attribute of being a tunable percussion instrument. Again, a mute is usually defined in terms of the transformational attribute of effecting a change in tone color (Flavell, 1970).

Rule Differences Not only do concepts have specific attributes which are involved in the public definition; there is also a rule governing the relationships between the attributes. Concepts differ in the type of rule involved in their definition. Bruner (1956) has indicated that one of three rules is usually involved in the definition of a concept. These are conjunctive, disjunctive, and relational rules. To these must be added the affirmation rule (Bourne, 1966).

The affirmation rule is simply stating the presence or non-presence of a single, relevant attribute. The type of concept in such circumstances is a primitive type since it has only one attribute which cannot be reduced to more elementary components. For example, timbre is a tonal element which includes a number of primitive concepts such as the concept of flute tone, saxophone tone, and the like. The affirmation rule is merely stating that such-and-such a tone is or is not the sound of a flute (or any other timbre).

A conjunctive rule states that two or more attributes must be present to constitute an instance of the concept. A simple example of conjunction is found in the concept of a "red square." Here the two attributes "red" and "square" are combined by the implied "and." It should be noted that both "red" and "square" are primitive concepts as well as defining attributes of the conjunctive concept. In music, an example of the conjunctive concept is "F sharp." The two attributes, "F" and "sharp," are combined to form a conjunctive concept. Learning a conjunctive involves learning the rule for combining attributes. In some cases, where the attributes (primitive concepts) are not known, it also means learning these as well as the rule. A large amount of conceptual learning beyond the primary level is the reorganization of primitive concepts into new combinations. Gangé (1966) has referred to such reorganization of primitive concepts by means of rules as "principle learning." He prefers to restrict concept learning to the more primitive level. Such a change in terminology clearly differentiates the process of identifying relevant attributes from the process of forming and using rules. Were Gangé's terms to be used, many of the concepts in the MENC book, Music in the Elementary School - A Conceptual Approach (Gary, 1967), would turn out to be principles instead of conjunctive concepts.

A disjunctive conceptual rule is one in which the presence of either one or another attribute suffices to define the concept. Red or square is a disjunctive concept by means of which stimulus objects may be classified. A "strike" in baseball is either called, or a missed ball, or a foul ball. A musical example of a disjunctive concept is the key signature as an indicator of mode. Thus, three flats indicate the major key of E^b or the minor key of c.

A third rule for combining attributes is a relational rule. In relational concepts, the attributes have a certain relationship to each other. Such concepts as "smaller," "farther," "higher," and "louder" are of the relational type. Take the concept "higher." In this case, two pitches may be compared; it makes no difference what the pitches are or how far distant they are from each other, the important aspect is the relationship of the first pitch to the second. Relational concepts are sometimes confusing to children who often consider a stimulus in absolute terms rather than in terms of its relationship to something else (CRM, 1971). For example, middle C may be a higher note in one situation but a lower note in another.

There is no need to provide an exhaustive listing of all of the possible rules for combining attributes. The point is that in all concepts some kind of rule is present which is independent of the attributes, and that learning a concept may not only involve identifying the relevant attributes but also formulating a rule by which the attributes are combined (Bourne, 1966).

Differences in Abstractness A concept is said to be concrete to the extent that its instances are tangible objects. A clarinet, drum, and recorder are concrete objects, whereas creativity, musicianship, a cappella ideal are concepts without particular, concrete referents. As concepts become more abstract, they usually consist of other concepts which define them (Flavell, 1970).

Differences in Generality Concepts may be arranged in a sequence from particular or limited categories to broader, more comprehensive categories. The more general concepts are sometimes known as superordinate concepts (CRM, 1971). For instance, apples, pears, and peaches belong to the category of fruit which, in turn, belongs to the category of food. In like manner, a Bach trumpet is a sub-class of the class of trumpets which in turn belongs to the brass family and are musical instruments.

Differences in Precision A concept is precise to the extent that our present knowledge would allow us to agree upon an explicit set of attributes for distinguishing instances from non-instances. Precise concepts are whole note, g minor scale (melodic), and perfect authentic cadence. Less precise concepts are phrase, consonance, style, and masterpiece. Even though these latter concepts do not lend themselves to precise formulation, they are often used effectively in identifying instances of the concept. Our inability to state the exact nature of a concept does not mean that it is not present in some inarticulate form in our cognitive structure (Flavell, 1970).

Differences in Power A concept is powerful to the extent that it satisfies two conditions: (1) most people would agree that the concept is, for some reason, an important and central one, and (2) its attainment facilitates the formation of other concepts. The dimension of power is one of particular interest to developmental psychologists and educators, since it has to do with critical points of concept development. The child acquires a multitude of concepts as he grows, but not all of them point beyond themselves to further acquisitions. It is such powerful concepts whose attainment the developmentalist wants to understand and the educator to foster. For example, Edgar Dale (1971, p. 1) has said, "To master a field of subject matter is to learn its key concepts." Learning the powerful concepts is also advocated by Woodruff (1961, p. 102):

> Organizing the curriculum properly means . . . selecting those concepts which are vital to life and leaving the others out. Our present curriculum is still organized around topics. Under this system it is impossible to be sure we have identified the significant concepts, and avoided spending time on the relatively insignificant ones.

An effort to carry out Woodruff's injunction was realized in music education with the publication of The Study of Music in the Elementary School - A Conceptual Approach. This book has organized musical experiences around powerful concepts which delineate the structure of music.

There is considerable agreement concerning which concepts are powerful in music. In the book just cited, the concepts of melody, rhythm, harmony, dynamics, timbre, and tempo are emphasized. Oleta Benn (1958, p. 342) has called for the "development of concepts of musical sound itself, with all its attributes of pitch, intensity, and timbre. In addition, we must develop the ability to judge duration." Meyer (1967, p. 248) considered pitch and time as "primary, pattern-forming parameters; dynamics, timbre and mode of playing are dependent variables relative to each other as well as to pitch and duration." He goes on to point out that pitch and time are not only basic dimensions of Western music, but "pitch-time relationships are also primary categories of organization in theory, notation, and musical terminology of most non-Western cultures."

Although experts agree that certain concepts are powerful, this only satisfies the first criterion by which a concept may be judged to have power. The other criterion is the effect of one concept in facilitating the attainment of others. To more accurately estimate the power of a concept in these terms necessitates a program of research on the development of conceptual sequences and hierarchies in children.

2. Differences in the Development of Concepts.

If the subjective, individual differences in the formation and use of concepts are examined, additional variation between concepts will be found. Flavell refers to these kind of differences as a variation in the subject-concept relationship. The differences in this relationship are differences in validity, status, and accessibility (Flavell, 1970).

Differences in Validity A concept is valid to the extent that it approximates a public definition. The validity of children's concepts varies in two ways:

(1) The child's operating concept may deviate in that it is broader, or narrower, or distorted in some way. The meaning is more or less stable for the individual but idiosyncratic compared to others. For example, a child's concept of "dog" may be limited to the family pet or it may include all small animals. Similarly, a scale may be narrowly understood as a particular pattern such as the major descending version found in the first phrase of "Joy to the World" or the concept may include all ascending melodies.

(2) The child's concept may not only deviate from the standard accepted by others, it may also fluctuate in its meaning for the child depending on the situation or expectation of the child. For instance, the concept of "higher" may on one occasion be used with reference to pitch but on another with reference to volume.

Both types of departure from validity can be noted in the development of concepts in children (and also in much adult new learning). For instance, Piaget's studies indicated that the young child's concept of quantity tends to be both different from and more unstable than the adult standard. Quantity is likely to have a unidimensional rather than tridimensional meaning. For the child, "more" means "taller" only, or "wider" only. It also may refer to one dimension, such as height, in certain situations and another dimension in other situations. An understanding of how certain powerful concepts stabilize and converge on their public prototypes during childhood is a major objective in the study of concept development (Flavell).

Differences in Status The status of a concept varies with the kind of behavior which is acceptable as a conceptual response. It is unclear to say that a child has a concept unless the type of response is also specified because concepts can be "had" in different ways. The concept user might have no name for a given concept; or he might know its name without being able to characterize it further; or he might be highly articulate in describing its full meaning and function. Children are very likely to "have" concepts which they cannot name, let alone define and discuss, and yet they operate consistently in non-verbal conceptual tasks. The adult is usually expected not only to operate using concepts but also to be able to examine a particular concept as an object within his cognitive structure. To expect the same of a child is to be unaware of the varying ways in which conceptual behavior may be manifested (Flavell, 1970).

Differences in Accessibility Some concepts appear to have been formed and used by a child on some occasions and yet, on a given occasion, he may be unable to use his concept. A number of reasons for the inaccessibility of a concept may be given including competition from other stronger concepts, unfamiliar context of the problem, inappropriate response requirement, unclear instructions, or different levels of motivation. The fact that conceptual behavior may be manifest on one occasion and lacking on another is sufficient reason for the researcher or teacher to be cautious in his evaluation of the conceptual development of a child (Flavell, 1970).

Some Considerations for Research on Concept Development

The developmental changes in the validity of concepts have traditionally received the greatest research emphasis, but a full account of conceptual development ought to encompass changes in status and accessibility as well. An awareness of these latter variations in conceptual development should make the investigator more aware of the fact that the young child may be utterly inarticulate about his concept and that he may also have trouble in bringing it into active service in all but the most obvious and compelling situations. When a concept is newly minted and freshly attained, it is likely to be unformulated, precariously utilizable, and, therefore, difficult to diagnose (Flavell and Wohlwill, 1969).

For Bourne (1966), a full account of conceptual development would cover two basic types of behavior: (1) concept formation and, (2) concept utilization. These two types may be further analyzed into various subbehaviors as shown in Table 1.

TABLE 1

SCHEMATIC REPRESENTATION AND EXAMPLES OF TASKS INVOLVING CONCEPTUAL BEHAVIOR

	Type of Behavior	
	Formation	Utilization
Attribute	Perceptual discrimination, labeling	Concept identification, sorting tasks
Rule	Formation of learning sets, positive transfer across problems based on the same rule	Rule identification, problem solving

(Bourne, 1966, p. 19)

Flavell (1970) has suggested that a research program on the development of concepts in children needs to have a comprehensive view of conceptual behavior. Such a program would take into consideration the following areas:

(1) Exact developmental dating of the appearance of a certain concept. In particular, there is a need to know which concepts emerge in a regular sequence and which emerge concurrently with each other. "Such information about the temporal relations in appearance is in turn a necessary first step in determining the effect of one concept in facilitating the attainment of another concept" (Flavell, 1970, p. 1033).

(2) Precise evaluation of the child's developmental status is also necessary if assessment is to be made of the effects of experience and training on cognitive growth. The magnitude of training effects can be inferred only from information about the cognitive status before and after the training experience. The basic requirements for a good training study are exact specification of initial cognitive status, of the final status, and of the intervening inputs and processes. These requirements are extraordinarily difficult to meet in practice, and even the most careful and sophisticated of the available training experiments raise more questions than they answer.

(3) Concern with diagnostic problems has resulted in the creation of a number of new testing procedures. The fundamental difficulty appears to be one of diagnosing accurately, the different ways in which a child may

be said to "have" a concept. Braine (1968) has been a strong advocate of nonverbal methods and has tried to devise procedures which will be sensitive to the first manifestations of a given concept. Flavell and Wohlwill (1969) have recently suggested that Chomsky's competence-performance analysis of language behavior and development may also be useful in thinking about conceptual development.

Briefly stated, their argument is that any given concept may first emerge as part of the child's intellectual competency considerably earlier than normal testing procedures would indicate. At this point in the child's development, the concept, while genuinely in the system, is exceedingly fragile and difficult to elicit and highly vulnerable to blockage by innumerable factors such as memory and attentional problems, interfering perceptual and conceptual sets, and the like. In the ensuing years, the concept slowly frees itself from the above performance limitations, gradually becomes consolidated, stabilized, and generalized, and eventually emerges as a reliably excitable cognitive tool in most appropriate situations and under most conditions of testing.

Flavell (1970, p. 1033) concludes, "If this general view of how conceptual development typically proceeds is even approximately correct, a great deal of painstaking research awaits us, because we know almost nothing as yet about the details of the process. . . ."

II. THE INTERRELATIONSHIP BETWEEN PERCEPTUAL AND CONCEPTUAL DEVELOPMENT

The relationship between perception and conceptualization is a problem which has intrigued philosophers and psychologists alike. It was the cause of the battle between the empiricists and realist philosophers (Boring, 1950) and it is at the root of much of Piaget's research program (Flavell, 1963). Phenix (1958, p. 305) has stated that "it is not only true that percepts underlie concepts; it is also true that concepts influence perception." Such a statement merely indicates that some relationship exists; unfortunately, it leaves unanswered the question of how concepts form from percepts in the first place or how, after concepts are formed, they influence subsequent perception. The problem is not one of interest only to the philosopher; it has important consequences for the teacher as well.

The procedure to be followed is to briefly consider the viewpoints of Gestalt psychology, of Bruner, Brunswik, Piaget, and Wohlwill. A concluding section will cover the influence of concepts on perception.

To clarify the discussion which follows, it may be helpful to view the development of concepts as a continuous line which begins with perception, leads to concepts, and reaches its final phase in linguistic competence.

PERCEPTION CONCEPTUALIZATION VERBALIZATION

Various Views On The Relationship of Perception to Concept Development

1. The Gestalt Position: One of the solutions to the problem used by the Gestalt psychologists was to make no qualitative distinction between perceiving and thinking. The model of perception was applied to thought processes with little or no adjustment. In Koffka's (1924, p. 49) view, "The ideational field depends most intimately upon the sensory, and any means that enable us to become independent of immediate perception are rooted in perception, and in truth, only lead us from one perception to another." What Koffka seems to say is that the "means" which permit the organism to go beyond the stimulus field are nothing but perceptions. These perceptions are of a higher, more veridical nature which have resulted from restructuring the stimulus field. How the field is restructured by perceptions themselves is not determined. Thus, the argument turns about itself in a circle without explaining how concepts are formed or how concepts may be the instruments whereby an immediate percept is formed. Another criticism of the Gestalt psychologists is their general avoidance of experimental work with children's cognitive processes (Wohlwill, 1968).

2. Bruner's Position: Unlike the Gestalt school which made perception equivalent to thinking, Bruner goes to the other extreme. For Bruner, perception is basically an inferential process in which the perceiver plays a maximal role in interpreting, categorizing, and transforming stimulus input. Unlike the Gestalt view of the perceiver who is essentially passive, Bruner's perceiver plays an active role in processing incoming information. Perception involves an act of categorization, says Bruner (1957, p. 123), and "the nature of the inference from cue to identity in perception is . . . in no sense different from other kinds of categorical inferences based on defining attributes." It seems clear that Bruner defines perception in essentially the same terms as others have defined conceptualization. For example, Harvey, Hunt, and Schroder speak of conceptual activity as follows (1961, p. 1):

> We assume that an individual interacts with his environment by breaking it down and organizing it into meaningful patterns congruent with his own needs and psychological make-up. As a result of this inter-change, perceptual and behavioral constancies develop which stem from the individual's standardized, evaluative predilections toward differentiated aspects of his external world. We will refer to such evaluation tendencies as concepts.

Bruner's model presupposes an adult perceiver whose conceptual categories have become more or less permanent. According to Wohlwill (1968), it would be difficult to apply Bruner's theory to the perceptions of a very young child who has neither perceptual constancies or stable concepts.

3. Brunswik's Position: Brunswik has emphasized the differences between perception and conceptual thought, rather than attempting to explain one in terms of the other. Perception, according to Brunswik, is a process by means of which an approximate, generalized picture of the stimulus pattern emerges. This is contrasted with the more machine-like process of reasoning which is more focused and precise. He says (Brunswik, 1956, p. 91), "The entire pattern of the reasoning process resembles the switching of trains at a multiple junction with each of the possible courses being well organized and of machine-like precision, yet leading to drastically different destinations." Although perceptual processes are only approximate and global, nevertheless they are an effective safeguard against a drastic error of adaptive behavior since they are closely bound to the objective world. Conceptual processes, on the other hand, when correctly used, lead to precise solutions, but otherwise may lead to counter-productive behavior.

4. Piaget's Position: Piaget believes that developmental stages exist in the realm of intellectual growth, but not in perceptual development (Wohlwill, 1968). He attempts to minimize the interrelatedness of perception and thinking. According to Piaget, the perception of the young child is "centered" in the sense that its organization is dominated by field effects. With age, the child's perception is progressively freed from its domination of field effects and becomes more logical in form. Like Brunswik, Piaget has noted the probalistic, approximate nature of perceptual judgments as opposed to the precise and certain results achieved through conceptual thought (Piaget, 1969).

Piaget considers perception and thinking as following two separate courses. On the one hand, conceptual growth leads to conservation and operational thinking; on the other, perceptual processes result in the perceptual constancies. In perception, the child scans the whole stimulus field for information. The only developmental change is in the extent and efficiency of this exploration. Piaget considers such a change as merely quantitative, not qualitative (Piaget, 1969).

> We are now in a position to appreciate the reasons that probably motivated Piaget's denial of the existence of stages in perception, while affirming it for mental development. . . He evidently believes that no meaningful criteria can be found in the area of quantitative perceptual judgments for distinguishing among different stages (Wohlwill, 1968, p. 480).

5. Wohlwill's Position: Wohlwill (1968) views perception and conceptual thought as two distinct processes which interact with each other. He states (p. 482),

> At a certain age level (in middle childhood) there is clear evidence of a 'perceptual compromise' showing the mutual interaction, rather than absolute separation, between perception and thinking.

Further support for Wohlwill's view is found in the earlier distinction which was made between concrete and abstract concepts. Such differences between concepts imply differential interaction with perceptual judgments. This interdependence between perception and thinking is the major premise for Wohlwill's conception of intellectual development. This conception is "built around the person's dependence on various aspects of information contained in the stimulus field" (p. 483).

The "various aspects of information" have to do with redundancy of information, selectivity, and contiguity.

(1) Redundancy: As a child proceeds from perception to conceptual thought, the amount of redundant information required decreases.

(2) Selectivity: As one proceeds from perception to conceptual thought, the amount of irrelevant information that can be tolerated without affecting the response increases.

(3) Contiguity: As one proceeds from perception to conceptual thought, the spatial and/or temporal separation of positive and negative instances can be increased (Wohlwill, 1968).

Let us consider these three dimensions of information processing in more detail.

(1) The dimension of redundancy: Redundancy may be understood as any surplus information which aids a person in classifying stimulus patterns. For example, in a sorting problem which involves the isolation of large, red marbles from small, green cubes, the dimensions on which these objects may be classified are size, color, and shape. However, only one attribute, say color, is actually needed to correctly sort the objects. The young child typically requires two or more dimensions in order to correctly categorize. The extra attributes are redundant. A musical example of redundancy in a problem solving situation occurs when a child identifies a major scale only which it is ascending, employs a particular rhythm pattern, and is played on the piano. Of course, for the child, the attributes of ascending pattern, rhythmic values, and timbre are not redundant because all are necessary for correct classification. Only after he learns that the pattern of pitch relationships determines the major scale do the other attributes become redundant.

A developmental trend in the direction of decreasing reliance on redundancy can be found in a variety of contexts. The clearest example of this comes from studies on the identification of geometric or familiar object stimuli on the basis of partial cues. The younger the child is, the more complete the figure need be; as he grows older, the less complete the figure need be for the child to identify it. It seems justifiable to regard such a task as becoming increasingly inferential as the amount of information which the subject has to "fill in" increases (Wohlwill, 1968). Bruner (1957) has postulated that the amount of redundant information required for a child to accurately identify an object is inversely proportional to the accessibility of that particular conceptual category. In other words, the more accessible a particular concept is, the less redundant information is needed to elicit the concept.

(2) The dimension of selectivity: Perception typically involves reception of sensory input of various kinds, i.e. color, size, shape; or pitch, duration, volume, timbre, and so on. At the perceptual level there is little, if any, organization of this input into relevant and irrelevant information. On the other hand, conceptual thought clearly involves selective attention and abstraction of relevant from irrelevant information. Wohlwill (1968, p. 485) remarks that

> It is thus noteworthy that one of the major developmental changes that seems to take place in the development of abstract concepts is precisely the differentiation of relevant from irrelevant, but more readily discriminable, attributes. This development is shown in various studies of concept formation; it may also lie at the heart of a problem which Piaget has studied intensively - the development of conservation.

(3) The dimension of contiguity: It is characteristic of conceptual processes that they enable the individual to deal with stimulus information whose components are widely separated in space or time. This is particularly significant in music where conceptual groupings inevitably contain instances which are separated in time. On the other hand, perceptual judgment requires that the stimulus objects or events be contiguous to each other in either space or time in order for them to be compared (Wohlwill, 1968).

Wohlwill has not indicated what internal mechanisms are responsible for these changes from perception to conceptual behavior. Instead, he has suggested what kinds of behavioral changes one might observe in diagnosing the formation of cognitive structures.

Conceptual Processes and Their Influence on Perception

The other side of the coin is the problem of how concepts may influence that which is perceived. A number of educators, musicians, and psychologists have offered their views on this problem. Philip Phenix (1958, p. 305) in his book, Philosophy of Education, declared, "What one observes with his senses is conditioned by the entire conceptual scheme. . ." Certainly auditory perception is included since Phenix later observes that a musician hears a piece of music in a more meaningful way than the untutored because he has a "scheme of musical concepts which direct him to listen for certain things which the layman would fail to notice" (p. 305). Leonard Meyer (1967, p. 216) has commented: "Any particular reality is a construct for there can be no perception without conceptualization, and every conceptualization entails abstraction from the particulars of a concrete object." The active role of the subject in organizing and constructing his inner world is emphasized by Neisser (1967, p. 3) in his comment that "whatever we know about reality has been mediated, not by organs of sense but by complex systems which interpret and re-interpret sensory information."

If the previous views are true, there is an obvious significance in them for teaching music since they suggest that learning concepts of music is the best, and perhaps the only, way of perceiving musical elements, their relationships, and meaning. Of course, such a conclusion is nothing new since music educators such as Aronoff (1969), Leonard (1959), and the MENC Committee on Elementary Education (Gary, 1967) have already developed programs based on this assumption.

Although the notion that concepts orient and regulate perceptual processing of sensory inputs is generally accepted, the method by which this is accomplished is not clearly understood. Such expressions as hypotheses testing, learning set, expectation, education of attention, and analysis-by-synthesis indicate the variety of approaches which have been advanced to account for the interaction of the individual's conceptual system with his environment. For the most part, the theoretical approaches mentioned above picture the organism as an "active" learner. Thus, Bourne (1966, p. 36) comments on how such theories view the learner:

> He is assumed to possess some selectivity. He operates in important ways on his environment. First, he may not respond to all available stimulus features, but rather select and attend to only certain aspects which, on the basis of a hypothesis, are considered relevant. Second, the subject decides upon and executes a response in conformance with his hypothesis which serves as a test of its adequacy.

1. Bruner's Approach: Bruner, Goodnow, and Austin (1956) reported a thorough theoretical and empirical analysis of many aspects of human conceptual behavior. They describe the process of forming and utilizing concepts as a series of decisions. The subject is presumed to begin any problem by deciding on some tentative hypothesis which may attribute importance (relevance) to one, some, or all of the dimensions which vary within a stimulus display. For example, a hypothesis may be that all

large, red, squares belong to the positive class. Thereafter, if the hypothesis on one trial proves correct, subsequent sorting is simply a matter of applying the rule. If the hypothesis is wrong, the learner must decide how to change it. Even if the hypothesis has worked on a given occasion, the subject may decide to reduce or augment its scope, i.e. drop or add some attribute. Thus, conceptual behavior is looked upon as a sequential decision-making activity in which each decision is contingent on an earlier one.

In the experiments by Bruner, et al. subjects followed definite strategies or plans of attack on conceptual problems. One strategy which was used frequently was called conservative focusing. This strategy is a systematic elimination of each of the possible hypotheses until the correct one is discovered. At each stage of his decision-making, the subject is left with a hypothesis which represents all that he has learned about relevant and irrelevant attributes up to that point. In some respects, it can be likened to an efficient method of playing the game Twenty Questions. Skilled players will systematically reduce the number of alternate possibilities to a small number and finally to the specific instance. Such a strategy might be followed almost unconsciously by a musician in trying to categorize an unfamiliar piece of music. Thus, he might first try to determine the historical period it belonged to by considering instrumentation, form, and stylistic characteristics. He may even go so far as to be able to suggest a composer by using his concepts of various composer's styles. It is obvious that this strategy is usable only by subjects with some training and experience with the problem situation (Bruner, et al., 1956).

Other subjects, particularly the naive and inexperienced ones, did not follow any particular strategy or used strategies which were less efficient. These latter strategies Bruner refers to as "scanning" and "gambling" (Bruner, et al., 1956, p. 84). The strategy which a person uses on one occasion may not be used in other situations.

Which strategy is used

> depends on a variety of factors determined by the problem, and conditions under which it is solved, and are subject to continual modification. Presumably . . . learning is an important factor in strategic behavior (Bourne, 1966, p. 39).

Unfortunately, Bruner does not provide an answer to the question of how such strategies are learned in the first place.

2. Harlow's Approach: One reasonable lead to answering the above question comes from the studies of learning set formation (Harlow, 1949, 1959). The basic experimental model for learning set theory is the "oddity" problem. In this task a subject is confronted with a set of three or more stimulus objects, one of which differs in some way from the rest. Suppose the stimulus array consisted of one square block and

two round ones. The subject would be allowed to choose one of the objects. If he chooses the "correct" one, he is rewarded. If he chooses incorrectly, all objects are removed and no reward is given. The problem is then repeated as often as necessary until the subject manifests a reasonable consistency in his response accuracy.

A new problem is begun soon after the last trial on the first problem. In the second problem, the relevant dimension may be size. In this case, the "odd" object would be a block which is smaller than the others.

Experiments with "oddity" problems have shown that there is a gradual increase in the performance accuracy within any one problem. A more important outcome is that on each successive problem, the solution is arrived at more quickly. A learning set has been formed in which subjects learned the oddity principle or rule. While stimulus attributes provided the cue, only knowledge of the rule can produce 100% correct choice on the first trial of each new problem (Bourne, 1966).

The oddity rule is a simple one, much more simple than Bruner's hypothesis testing strategies, and yet experiments with children suggest that it is not obvious for the young or naive child. It must be learned, and it is used without error only after extensive practice. The present research has some relationship to the "oddity" experiments.

3. Neisser's Approach: Neisser, like the others, thinks of an active learner who "constructs" the environment in ways which are related to his conceptual system. He says (1967, p. 10),

> The constructive processes are assumed to have two stages, one of which is fast, crude, holistic, and parallel, while the second is deliberate, attentive, detailed, and sequential.

Two points should be made in connection with Neisser's approach. First, his theory is somewhat parallel to computer information processing systems. Such an approach is quite compatible with Wohlwill's suggestion that a cybernetic model of cognitive activity might provide a useful approach in leading to a better understanding of the relationship between perception and conceptual behavior. Wohlwill suggested that "the operation of scanning mechanisms as characteristic of perception, as against digital mechanisms intervening in reasoning. The process of developmental change could then be conceptualized in terms of varying forms of interaction between these two" (Wohlwill, 1968, p. 488). The second point is that Neisser's two-stage model bears an obvious relationship to the positions of Brunswik and Piaget which were earlier described in which perception is characterized as holistic and approximate and conceptualization as precise and detailed.

The importance of Neisser's theory of "feature analyzers" lies in the emphasis he gives to auditory perception and its implication for musical perception. With regard to listening he has said (p. 194),

> Listening is a constructive process. . . It has both a passive and active mode. A preliminary analysis, made by a relatively passive, preattentive stage provides information which guides the more active process of synthesis itself. . . The Listener must have a set of rules, phonetic, syntactic, or what you will. It is the employment of these rules that makes analysis-by-synthesis more powerful than such methods as correlation or filtering. These rules are organized as a strategy for selecting the order in which patterns are synthesized.

Further on he states (p. 194),

> One makes a hypothesis about the original message, applies rules to determine what the input would be like if the hypothesis were true, and checks to see whether the input is really like that.

Neisser's hypotheses, rules, and strategies appear to have much in common with the viewpoint of Bruner. Unfortunately, Neisser has not shown interest in the developmental aspects of his theory.

4. Other approaches: There are other theoretical approaches which have been developed to account for some aspect of the relationship between perception and conception. Broadbent (1958) has developed a "filter theory" to account for a preliminary screening and evaluation of sensory input. The filter process is under the control of higher cognitive processes. Kagan and Kogan (1970) have explored the effect of attention and motivation on perception and conceptual behavior.

The relationship between perception and conceptual behavior is indeed complex. It may be noted that each of the theoretical approaches which have been mentioned interpose one or more intervening variables between what is perceived and what is conceived. Although the precise mechanisms which mediate between the perceptual and conceptual systems are not known, the basic tenet that perception underlies conceptual development and that concepts effect how we perceive is generally accepted.

III. THE RELATIONSHIP BETWEEN LANGUAGE AND CONCEPTUAL DEVELOPMENT

The interplay of language and human conceptual behavior, though exceedingly complex, is a fundamental problem; it pervades most, if not all, of the theory and research of conceptual development (Bourne, 1966). The study of verbal learning is an expansive enterprise which has occupied psychologists for many years. Often this work has overlapped research on conceptual behavior. In his book, Language and Thought, John Carroll (1964, p. vii) writes:

> I feel certain that the psychology of language and thought has important implications for particular problems in education, and for everyday life.

Asahell Woodruff (1970, p. 53) has expressed a similar view in his comments concerning the problem of the role of language in conceptual behavior. He says:

> Educators face the responsibility of resolving this problem. Its resolution is of central importance to the solution of instructional patterns.

Granted that the problem is an important one; however, it is by no means one which lends itself to easy solution. Flavel (1970, p. 1041) pointed out that it has long been assumed that language and conceptual development must interact in some way but that

> no one has yet formulated the possible nature of this interaction in any really clear and precise way, particularly with respect to putative influences of language acquisition on conceptual growth.

Piaget also admits the difficulty of the problem and declares it to be one of the most controversial issues in psychology. He stated (1970, p. 92),

> Any serious answer to the question of how linguistic and logical structures are related must, of course, be provisional. We cannot obviously solve the problem here; all we mean to do is to indicate what, from the structuralist perspective and taking recent developments in linguistics into account, the state of the question is.

The controversy involving language and conceptual behavior can be reduced to two principle questions: One, to what extent does language determine the nature of the concepts we have; and, two, what is the acceptable and appropriate way of measuring conceptual behavior?

Different Views on the Effect of Language on the Development of Concepts

In considering the relation of language to conceptual thought, B. L. Whorf has developed a hypothesis, known as the Whorfian hypothesis or linguistic-relativity hypothesis, which affirms that language determines the mode of thinking. In Whorf's words (1956, p. 212 ff):

> We cut nature up, organize it into concepts, and ascribe significance as we do, largely

> because we are parties to an agreement to organize it in this way - an agreement that holds through our speech community and is codified in the patterns of our language. The agreement is, of course, an implied and unstated one, but its terms are absolutely obligatory; we cannot talk at all except by subscribing to the organization and classification of data which the agreement decrees.

Bruner, although not in full accord with Whorf's hypothesis, suggests that language learning is one of the tools of cognitive advancement; however, Piaget and some of his major collaborators suggest that language develops as a result of the advancement of cognitive structures (CRM, 1971). For example, Furth (1966) strongly suggested that the acquisition of a linguistic system cannot be a necessary condition for the growth of human thinking, at least up to the level of concrete operations. This point is reinforced by Piaget (1970, p. 94):

> Even sensori-motor intelligence already involves certain definite structures which derive from the activity of coordination and are prior to rather than derivative of language.

Both of the above views have specified that below a certain age, intellectual structures develop independently of language; however, after the stage of concrete operations is reached (about seven years of age), Piaget recognizes the effect of speech in facilitating or partially structuring further conceptual development.

Woodruff apparently considers language to have little, if any effect on the development of concepts. He says (1970, p. 8),

> In spite of the noisy and pervasive function of speech, the primary process of storing and using experiential memories is basically a non-verbal one. This statement will provoke objections from many people, chiefly those who believe concepts as essentially verbal in nature.

Elsewhere (1966, p. 220) he explains that

> Concepts are acquired only through the basic senses, when they are in direct contact with real things. Language is not a means for getting such knowledge, and verbal activity is not equivalent to basic sensory learning. Language is a means for referring to our concepts, and we can use it to assist ourselves in getting our concepts organized and clarified after the senses have supplied us with our basic cognitive data.

For Woodruff the place of language is adjunctive to conceptual behavior, not central.

A good portion of the controversy concerning the effects of language on the development of concepts stems from the failure to specify (1) which kind of concepts are or are not facilitated by verbalization and (2) which age level is being considered. If the development of primary concepts at the primary level is being considered, few would argue in defense of verbal learning; however, at a later stage, where basic concepts have been acquired, verbalization may be necessary for the development of more abstract concepts. Such a two-stage development is considered by David Ausubel (1965, p. 96 ff):

> Children (below the age of twelve) are closely restricted to basic empirical data in the kinds of logical operations they can relate to cognitive structure. Thus, in performing "class inclusive and relational operations," they generally require direct experience with the actual diverse instances underlying a concept or generalization as well as proximate, non-verbal contact with the objects or situations involved. During the elementary years directly presented and verbal materials are too distantly removed from empirical experience to be relatable to cognitive structure.
>
> Beginning in the junior high school period, however, and becoming increasingly true thereafter, prior empirical and non-verbal experience is no longer essential before concepts and generalizations become potentially meaningful. It is true, of course, that the pupil's established verbal concepts must have been preceded sometime in the past by direct, non-verbal experience with the data from which they were abstracted; but once these concepts are sufficiently well consolidated . . . new learning material is logically relatable to cognitive structure without any direct or non-verbal current reference to empirical data. The concepts and generalizations of the adolescent, therefore, tend more to be second-order constructs derived from relationships between previously established verbal abstractions already one step removed from the data itself.

Carroll (1964) has further delineated concept development as consisting of three stages. The first stage is concept formation. Carroll believes that the first kind of concepts formed are of the primitive type and specific to tangible objects in the child's environment such as the family dog. The development of these primitive concepts can be accounted for, in Carroll's opinion, in terms of S-R theory. Obviously, the concept at this stage is non-verbal and highly individualized. Even so, evidence for the existence of the concept is gained by the child's similar response to the dog even though the animal appears in a variety of situations.

The second stage of development is associating a label or word with the object. Here, it is first necessary to consider the word as a physical entity apart from its meaning. That is, the word has certain phonetic qualities which characterize it without considering its meaning. These characteristics of the sound of the word must be learned in a variety of contexts. The child must learn that the word "dog," as spoken by different people with different inflections and the like, is the same word each time in spite of the auditory variation.

The third phase is learning the public meaning of the word. After the name (as a sound) becomes associated with the concept "dog," there is an interaction between the public, denotative, meaning of the word and the privately-held concept. Through an interaction which takes place by means of social reinforcement, the child's concept of dog gradually comes to have the approximate meaning of the conventional definition of the word (Carroll, 1964).

Several points should be added to the previous outline. First, the concept is most likely to have an affective component as well. This affective component is not changed later when the concept becomes attached to the word but becomes the connotative aspect of the word's meaning. Second, even after the concept and the word become joined, the two are not identical. Piaget (1970) insists that the concept's relation to the experiences from which it originated is always closer to these experiences than to the public definition of the word. Thirdly, the concept being learned in the above outline was considered to be of a primitive type; more abstract concepts do not necessarily require the direct, non-verbal experience.

Various Views on Verbal Capacity as a Measure of Concept Attainment

The question at the beginning of this chapter was "What is a concept?". The answer given at the time was basically in terms of non-verbal behavior. Now, it seems appropriate, to mention that this answer would not be acceptable to a number of researchers. Carroll has cautioned against interpreting experimental results without first understanding the experimenter's definition of conceptual behavior (1964, p. 82):

> You must be careful to note what definition of concept learning is being used in a given instance. One definition has the virtue of complete objectivity; according to it, a person has learned a concept when he can with a high degree of reliability discriminate between instances and non-instances. This definition is usually satisfactory, but many individuals who know a concept by this definition are not able to formulate the concept

> verbally or to communicate it to others. In fact, several experimerts have shown that it is possible to learn a concept without being aware of the basis for it. . . Because such "unconscious" concept formation is possible, in some contexts, it is useful to define concept learning in terms of the ability to recognize instances *and* the ability to formulate descriptions, or to construct instances of the concept.

The views of Bloomfield, E. B. Hunt, and Archer may be characterized as emphasizing verbal behavior as a primary measure of conceptual learning. Perhaps the most extreme position is that of Bloomfield who insists that the idea of concepts is completely reducible to that of a word's signification. In fact, Bloomfield (Piaget, 1970) believes that there are *no* concepts, that what is mistakenly called a concept is simply the signification of the word. In the same way, E. B. Hunt (1962) holds that conceptual behavior should be defined in terms of the learner's ability to use labels or words properly. Moreover, he would require the learner to state a rule or principle which was used in the solution to the problem. Conversely, a learner who was unable to justify his classificatory behavior verbally does not have a true concept according to Hunt. Archer (1964, p. 238) has placed verbal capacity at the heart of conceptual behavior. For him, concepts "are meaningful words which label classes of otherwise dissimilar stimuli." He goes on to point out that a shortcoming in the ability to use words is the single most important factor in the slow acquisition of concepts by preverbal human subjects and lower organisms. Bourne (1966) notes that researchers who emphasize the verbal response aspect of conceptual behavior are not really dealing with the question of how concepts are formed but, rather, with the question of how they are utilized.

For the developmental psychologist the more interesting and challenging problem is how concepts are formed in the first place. This usually means accepting the notion that concepts have a preverbal stage of development. As a consequence, non-verbal methods of studying conceptual behavior have been used to diagnose conceptual development. A strong advocate of non-verbal methods is Braine who has been critical of Piaget's clinical method for its heavy reliance on the verbal report of the child. He observes in reference to Piaget's theory (Braine, 1968, p. 172):

> No theory which postulates levels of conceptual development can be regarded as definitely established when the supporting data are obtained through extensive verbal communication with the subject.

The controversy between Braine and Piaget centers on the kind of response which is an acceptable demonstration of conceptual behavior. Using non-verbal measures, both Braine and Bruner have discovered children are able to show conservation at the age of four or five years rather than at ages seven or eight (Sigel, 1968). Smedslund has referred to the kinds of responses which children in the Braine and Bruner experiments exhibit as a "symptom response." He tends to discount the significance of such responses as not very helpful in understanding the development of cognitive structures (Sigel, 1968).

Be this as it may, Sigel points out that using verbal criteria to diagnose conceptual development is somewhat risky since the child may have difficulty in finding the right words which explain his thinking. Furthermore, the investigator's decision to accept a verbal answer as correct or not assumes that a fairly sharp line differentiates one verbal report from another. This is sometimes difficult to do and the decision by the experimenter can be sometimes arbitrary (Sigel, 1968).

Piaget himself seems to be open-minded as to the use of non-verbal methods for assessing conceptual growth. In his preface to the book by Laurendeau and Pinard (1962), Piaget states:

> Verbal thinking seems to me marginal to real thinking, which, even though verbalized, remains until about eleven or twelve years of age centered upon action (p. XII). . . verbal thinking, therefore, no longer seems sufficient for the investigation of the child's thinking; it provides a series of instructive indications, which must, however, be related to other findings derived from operational tests proper, (p. XIII).

More recently Piaget has advocated techniques such as those of Braine as useful in diagnosing conceptual status. He suggests two techniques: the first, he refers to as Braine's method of transformational analysis; the second, as an operational analysis of Inhelder, Sinclair, and Bovet. These two methods according to Piaget (1970, p. 94),

> enable us to analyze the correlation between syntactic and operational structures, at least at a particular point; we are even in a position to guess just where there is interaction between the two and which of the linguistic or logical structures are prior, which are posterior, in the process of construction.

The widespread use of the above methods is a matter for further research. "It will take considerable investigative ingenuity and talent if we expect to make significant progress in the near future" (Flavell, 1970, p. 1044).

In conclusion, one must agree with Phenix who says that a degree of modesty is in order in making claims for language, for language can be confining as well as liberating. "Powerful as words are, they need to be supplemented by other symbolic forms as music, the dance, painting, sculpture, and religious ritual" (Phenix, 1958, p. 413 ff).

CHAPTER III: DEVELOPMENT AND IMPLEMENTATION OF RESEARCH PROCEDURES

Prerequisite to the onset of the experiment an instrument for diagnosing music concept development needed to be developed. Two forms of the instrument were produced with the Long form serving as a prototype of the Short form. The auditory features of both forms were identical; however, the task design varied between the two forms. The details of each form are described separately in this chapter as are the administrative procedures, the sample, and research design.

I. DEVELOPMENT OF THE LONG FORM OF THE INSTRUMENT

The Tasks as Instruments for Eliciting Conceptual Behavior

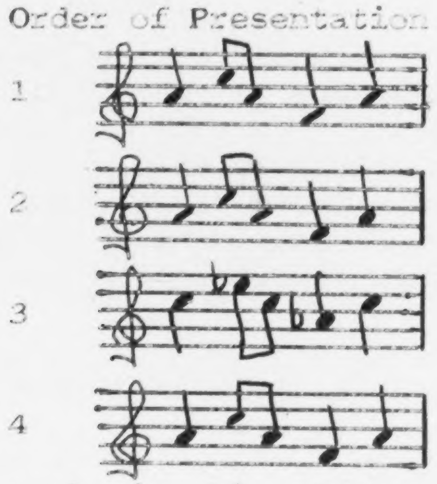
An initial problem was to develop an instrument which would reliably measure valid conceptual behavior. Several researchers had developed a task design which used a sorting strategy as a means of measuring concept attainment (Bourne, 1966; Bruner, 1956). With the Long form a subject was required to perform a task in which he was to sort melodic patterns into classes. Four patterns were heard by the subject who had been instructed to find three patterns which "go together" and one pattern which was "different". Using this strategy, a child was required to attend to specific features of the melodic pattern in order to find its relationship to the other patterns. The problem was to discover which feature was relevant to categorizing each pattern.

A melody contains several features such as pitch, tempo, timbre, durational values, and volume. Ordinarily these features vary freely from one melody to another. In these tasks, however, only two features were allowed to vary while the others were held constant. The variations in the two features across the four melodic patterns may be represented as r1, r1, r1, r2 and i1, i2, i3, and i4. The "r" designates the relevant feature since it is essential to the sorting task as can be seen by the fact that the first three patterns can be grouped together since all r's are held constant. The only variation of the relevant feature was in the fourth pattern. Those instances in which the relevant feature is unvaried are positive instances; the instance in which the relevant feature varies is the negative instance. The "i" designates the irrelevant feature. Since this feature varies with each instance, it is irrelevant to solving the sorting problem. The subject was asked to place the positive instances together and isolate the negative instance, i.e., the "different"

FILMED FROM BEST AVAILABLE COPY

one. Figure 1 shows a notated example in which the relevant feature is pitch. To identify the negative instance required the subject to use the concept 'higher'.

Fig. 1. Concept task where pitch is relevant



The design of the tasks in this study was similar to conceptual tasks as described by Sigel (1964), Gangé (1965), Carlsen (1969) and Aronoff (1969). Guilford denotes this type of sorting problem as an exclusion type task (Guilford, 1966).

Some Principles Involved in the Design of the Tasks

In order to say that tasks elicit conceptual behavior, certain psychological principles must be incorporated into the design of the tasks. First, it must be clear that the behavior in question could not have arisen as a simple S-R connection, or as a chain of connections. An S-R connection might result if one used only a single task and repeated it many times without varying the position of the positive and negative instances as in PPPN, PPPN...PPPN. A subject would likely choose the N instance after successive trials without necessarily displaying conceptual behavior.

To say that behavior is conceptual, it must be equivalent across a series of problems which contain some novel features. Therefore, tasks were constructed which contained a common, relevant feature from task to task but varied in their melodic patterns. If a child responded correctly over a sequence of tasks which were novel in some way, his behavior was taken as a demonstration of concept utilization (Gangé, 1965).

Inherent to the design of the instrument was the notion that consistency is a prime manifestation of conceptual behavior. Thus, the tasks within a sequence were homogenous. Furthermore, the approach to scoring required that some level of consistency be achieved as a criterion for concept attainment. With the Long form the scoring procedure used both a consistency criterion and an accuracy level. The consistency criterion was arbitrarily set as five consecutive correct responses. Since there were 20 tasks in each sequence, a subject might reach the criterion anywhere between the fifth and twentieth task. The point in a sequence where the criterion was reached was the consistency score. When a child reached the consistency criterion, he was excused from further work on that sequence. If a child reached the twentieth task of a sequence without achieving the criterion, he was excused and his failure to reach the criterion was recorded.

An accuracy score was simply the total number of correct responses a child made on a sequence. Both scores were needed to guage conceptual behavior more accurately. To illustrate this point, suppose that the following two sets of responses were obtained.

Subject A +0+0+ 0+0+0 +0+0+ 0+0+0

Subject B 0+000 +00+0 00++0 +++++

Both subject A and B had the same accuracy scores; however, Subject A did not reach the consistency criterion, whereas Subject B, after an uncertain start, appears to have discovered the solution to the problem as indicated by the consistency score.

A second principle of conceptual behavior is that the correct response could not be achieved on the basis of some extraneous cue. Sources of invalidity may sometimes arise from the measuring instrument or the testing environment. Faulty design of the instrument would be evident if the correct answers were determinable on the basis of some extrinsic aspect of the problem rather than intrinsic features of the melodic patterns. For example a child would be able to correctly respond if the negative instances of the concept appeared in the same position in successive tasks, as 1-PPPN, 2-PPPN, 3-PPPN, simply on the basis of the order. To disallow such a probability, the negative instance was randomly assigned to one of the four order positions, i.e., first, second, third, fourth.

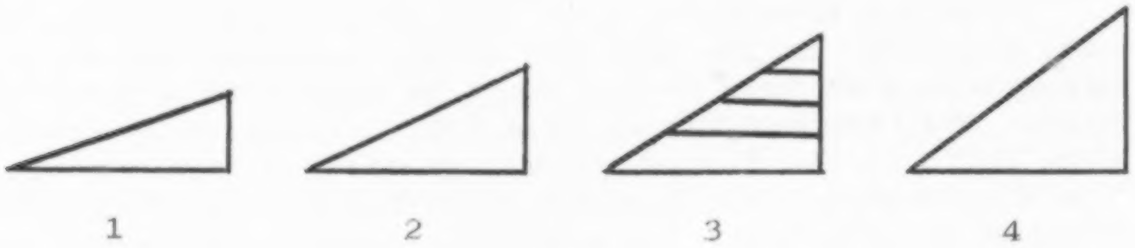
Although some other sorting tasks have used three instances (Carlsen, 1969), the tasks of the Long form were organized as a set of four instances because there was more flexibility in locating the negative exemplar with four instances. Of course, even more instances would mean greater flexibility but another problem would develop. Since auditory patterns must be presented successively rather than simultaneously, as is sometimes true in visual sorting tasks, any increase beyond four instances would have added an undue strain on the memory. Since these tasks were not concerned with the effect of memory on conceptual behavior, four instances were considered appropriate to the purposes of the study (Guilford, 1966; Bourne, 1966).

A third principle used in designing the tasks was that the task must require abstract thought in the form of a generalization growing out of experience with specific instances. To illustrate the principle, suppose the child is given the task of sorting toy vehicles into appropriate groups. If three tractors and one fire truck were in the collection of toys, it is plausible that the child would set the fire truck aside. Now, if all three tractors were identical in all their features, i.e., color, size, shape, etc., and the fire truck varied in these features, the conceptual problem of grouping three identical tractors would not be very significant. Indeed, some theorists view such behavior as a form of multiple discrimination learning rather than conceptual behavior (Gangé, 1966). Admittedly, the difference between conceptual thought at its lower levels and S-R learning at its higher levels is a subtle one and psychologists are not unanimous as to the proper line of demarcation (Melton, 1964). To return to the toy vehicles, suppose the tractors were not identical but varied in one or more features, say color and size, yet their shape was constant. The solution to the sorting problem would then be based on abstracting the shape feature from the other variations and using it as the basis of categorization. Since the fire truck did not possess the same shape, it ought to be excluded.

To restate the principle in a slightly different way, the positive instance of a class may vary in one or more features, but on at least one feature there is no variation. The negative instance would not display this feature, say pitch, at the same value. The fact that there is a variation in another feature means that the subject cannot group the instances on the basis of multiple common features but must recognize only one common feature as relevant to the solution and disregard other varying features as irrelevant. This type of conceptual sorting task was described by Sigel and was used as a model. Sigel (1964, p. 241) commented that "to teach a concept, juxtaposition of two competing concepts will force a child to reflect and think, rather than respond with what he already knows." The assumption was that variation of two features, one relevant, the other irrelevant, would involve a conceptual conflict which could be satisfactorily resolved by using the "rules of the game." Figure 2 shows visually the nature of conceptual conflict.

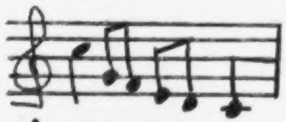
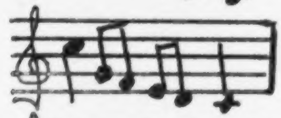
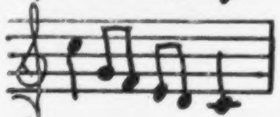
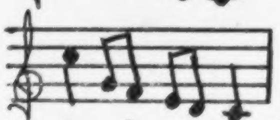
Fig. 2. Visual model of concept task

Find three shapes which go together and one which is different.



To translate the model above into an auditory task required that two features be systematically varied, one being relevant, the other irrelevant, with all other features remaining constant. An example of such a task is shown in Figure 3 where the melodic pattern remains constant and tempo and volume vary. Volume varies across all four instances, whereas tempo varies in only one instance.

Fig. 3. Tempo and volume changes across melodic patterns

Order	Melodic Pattern Constant	Tempo Varied (Relevant)	Volume Varied (Irrelevant)
1st		84MM	Base level db.
2nd		84MM	+ 2 db.
3rd		112MM	+ 4 db.
4th		84MM	+ 6 db.

Tempo in the above task was designated as the relevant feature because it best fits the rules of the game. Therefore, volume was an irrelevant feature. Since the rules were to find three patterns which are the same in some way and one pattern which is different, instances one, two, and four ought to be grouped together since they all had the same tempo whereas the third instance had a faster tempo. To solve this problem a child had to use the concept "faster" in order to exclude the negative instance. In this task the variation in volume across all four instances would not allow volume to serve as the basis for classification since the volume level did not remain constant for three of the four instances as did tempo.

Summarizing, the three principles which influenced the design of the Long form tasks were (1) demonstration of consistency, (2) control of extraneous variables which might produce non-conceptual correct responses, and (3) inclusion of both relevant and irrelevant features.

Technology Involved in Presenting the Tasks

Previous experience with similar conceptual tasks which had been presented on the Bell and Howell Language Master led to the assumption that this device might be appropriate for this study also. The Language Master is a type of tape recorder-player which uses short strips of recording tape affixed to 3 x 7 inch cards. The advantages of the Language Master for this study were (1) simplicity of operation, (2) a melodic pattern could be recorded on a single card, (3) cards could be sorted into categories, (4) a card provided a visual referent.

In spite of these advantages the use of the Language Master had to be restricted. The principal shortcoming was that it did not reproduce the melodic patterns with satisfactory fidelity.

The sound system which was finally developed began with the Moog synthesizer, Model 3P, which was used to produce all the melodic patterns. Use of the Moog permitted more accurate quantification of the various features being manipulated in the tasks than other sources of sound which had been tried, i.e., electronic organ, bells, and trumpet. From the Moog the electronic signal was transmitted directly into an Ampex tape recorder, Model 860, from which a master tape was made. Duplicate tapes were then made on the Sony cassette tape duplicator at the University of Nebraska Audio Lab. Cassette tape recorders were used at each school to present the tasks to each child. The cassette tape was judged to faithfully present the melodic patterns and the variations in the features.

Unfortunately, this system lacked the visual and manipulative advantages of the Language Master. To regain these advantages, a tandem arrangement was used in which blank cards were run through the Language Master synchronously with the patterns as they were played on the cassette tape. The result was a kind of technical ventriloquism which combined the advantages of the cassette and the Language Master.

Specific Details of the Design of the Tasks

Selection of the Concepts to be Studied As has been mentioned, the tasks were designed to measure the attainment and use of the concepts of louder, faster, higher, and shorter. Of course, these concepts are components of volume, tempo, pitch, and duration. Why were these concepts selected rather than others? These concepts were selected because they were believed to be the type most appropriate for the age group being studied. Flavell (1970) has shown that concepts may be classified according to various types. The starting point is the primitive concept. Such a concept involves only one relevant feature (Bourne, 1966). A single color, such as yellow, may elicit the primitive concept of yellow. In like manner, the timbre of a drum or bell is such as to produce a primitive concept. Other primitive concepts are the music concepts of louder, faster, higher, and shorter. Of course, these concepts all have reciprocals, viz., softer, slower, lower, and longer. Primitive concepts arise from stimulus features that are so elemental that they are merely pointed to or affirmed as being present in the stimulus display. Even though children use concepts which are more complex than these primitive ones, it seemed that study of primitive concepts was a suitable point of departure (Pflederer, 1967;

Andrews, 1968). Another advantage of these particular primitive concepts was that they were intimately bound up with the basic elements of music (Gary, 1967).

Another reason for selecting these concepts was their objective, physical counterparts were susceptible to controlled manipulation and quantitative description. Volume changes were measurable in terms of decibels, tempo changes in metronome units, pitch changes in scale degrees with fixed frequencies, and duration in milliseconds.

The third reason for studying these concepts was that they apparently bear some relationship to each other. Stevens (1938) has shown that perception of pitch varies in relation to both the frequency and the volume of sound. The possibility of an interaction between pitch concepts and volume concepts was implied by Andrews (1968). Similarly, duration and tempo bear an obvious relation to each other in that they both are temporal qualities of sound. The development of the concepts of louder, faster, higher, and shorter may proceed in parallel or independent stages; in any case, research has not yet illuminated either the stages or the possible relationships.

Construction of the Melodic Patterns A sequence used twenty different melodies, one for each task. These melodic patterns were original melodies rather than familiar melodies which might have given some children an advantage due to previous experience with the melody.

The melodic patterns possessed these characteristics:

The patterns were melodic. They were intended to be heard as a tune, as a whole musical unit, and as a pattern rather than as isolated, random tones. Inasmuch as the auditory materials were short, they could be better classified as a motive or a pattern rather than as a complete phrase. The term "pattern" was deemed appropriate because it suggested an integral relationship between the pitches or a Gestalt.

Patterns were tonal. A tonal center was either present or implied. Patterns varied in key, mode, and use of accidentals.

Patterns were within the accepted singing range of the children being studied (Nye and Nye, 1970). The range was between middle C and its octave. Six of the patterns added a d to the range.

The length of the patterns was relatively short. Patterns contained four to five beats and lasted a total of three to four seconds. Within each pattern tones of varying length were used, the most common being the quarter and eighth note values. On occasion dotted quarters and eighths were included. Both duple and triple meters were present in the patterns.

Variations in the Patterns The patterns were varied within a task by the changing of the whole pattern rather than a portion of the pattern. Patterns were not deformed by changing one or two tones but were changed as a unit. For example, a pattern may have been repeated four times, three times it was in the key of D, the fourth time it was in the key of F. This procedure differed from that used by other researchers (Pflederer, 1964) who varied only certain tones of the melody. This procedure was not followed because it involved the problem of finding an acceptable point in the melody at which to make a change. Whether a deformation occurs at the beginning, middle, or end of a melody may be a significant element in conceptual performance (Farnsworth, 1969). In this study, whenever a feature was changed, the change affected the whole melody.

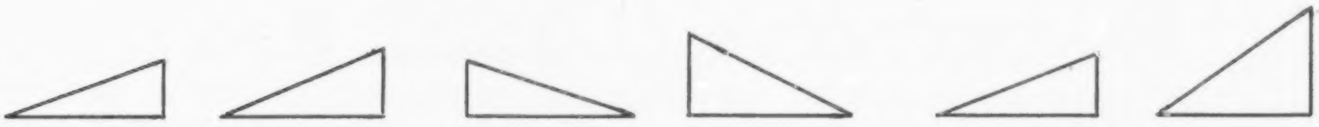
This holistic variation technique was readily applied to such features as volume and tempo. With pitch the pattern was varied by transposing upward a given interval. In the case of duration, the problem was to not confound tempo changes with durational changes. If all tones of a pattern were made proportionately shorter, the effect would be that the melody would seem to be at a faster tempo as well. If only one tone were shortened, the problem of choosing the best location for a change arises again. The solution was to shorten each tone, but not to change tempo. This resulted in a staccato sound. In other words, the durational variation of a pattern consisted of a variation between legato and staccato. The precise values of the tones in terms of duration will be indicated later.

The Values of the Relevant and Irrelevant Features

A major problem in developing the instrument was determining the value of the change of relevant and irrelevant features. The problem is inherent in the definition of a concept. The definition used in this study was "a concept is an internal structure whose existence is inferred whenever two or more discriminable objects or events have been grouped together and set apart from other objects or events on the basis of some common feature" (Bourne, 1966, p. 1). The difficulty here is with the word "discriminable." For objects or, as in this experiment, patterns to be distinguishable from each other, how large a difference must there be between the features? The problem of discrimination is illustrated in Figure 4.

Fig. 4. Sets of triangles illustrating varying degrees of discriminability

In which set of triangles are the shapes more discriminable?



Set A

Set B

Set C

Obviously, set C contains triangles which can be more readily discriminated than those in set A. But, what about the triangles in Set B? Is the difference between them sufficiently discriminable? To state the question in terms of an auditory problem, if three melodic patterns are presented at a given decibel level, by how many decibels should the fourth pattern be changed? A scale of values was needed for each feature which would indicate the acceptable range of discriminability before the tasks could be produced. The scale should be precise enough to indicate the lower level as well as the upper level of the acceptable range.

The problem of discriminability is not merely a practical problem related to the production of changes in the features of the tasks. It is an important theoretical issue since conceptual sorting behavior is certainly affected by the discriminable differences in the stimulus display (Flavell, 1970). Very small differences might not be detected and performance would be poor; however, a very large difference between two features may not require much in the way of conceptual thought. If a sorting problem consisted of grouping three peas and isolating a watermelon, one could justifiably question the reasonableness of such a gross difference to demonstrate abstract thinking.

The method used in developing a scale of discriminable values was to arrive at a subjective scale through estimates of changes as described by Stevens (1938). Four auditory discrimination measures were produced and administered to groups of children between the ages of six and ten years of age. The children were asked to offer estimates of difference in volume, tempo, pitch, and duration. A detailed description of the procedure used to assess the discriminability of volume changes follows.

Children were instructed that they would hear a short melody which would be repeated after a pause. When the melody was repeated, it would be either exactly the same or increased in volume. If a child did not hear any change, he was to make a zero on his paper, if there was an increase in volume, he should write 1, 2, or 3 depending on how large he estimated the change to be. He was to use 1 if he thought the change was small, 2 for a medium change, and 3 for a large change. The children were then given examples of the range of change which was used in the test. After these examples had been played and discussed, the children listened to 20 pairs of melodic patterns.

The values of the changes in volume from the base level were either zero, two, four or six decibels above the base level. These different values were distributed irregularly among the 20 items of the test. There were five items in which there was no change, five which changed by two decibels, five which changed by four decibels, and five which changed by six decibels. Each set of five used the same melody. Whenever a pattern was changed, it was always the second member of the pair which was increased in volume.

The results of this measure are shown in Figures 5 and 6.

Fig. 5. The % of children who offered differing estimates of the size of a change in volume when there was no change

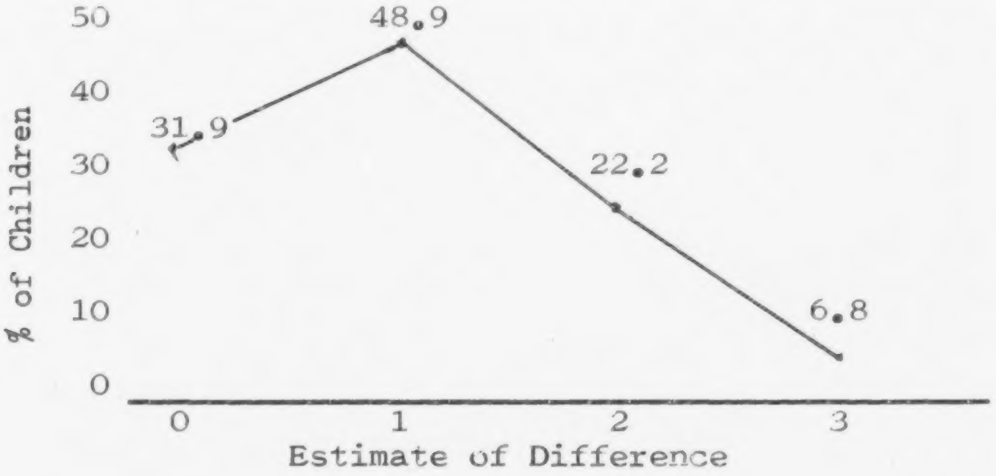
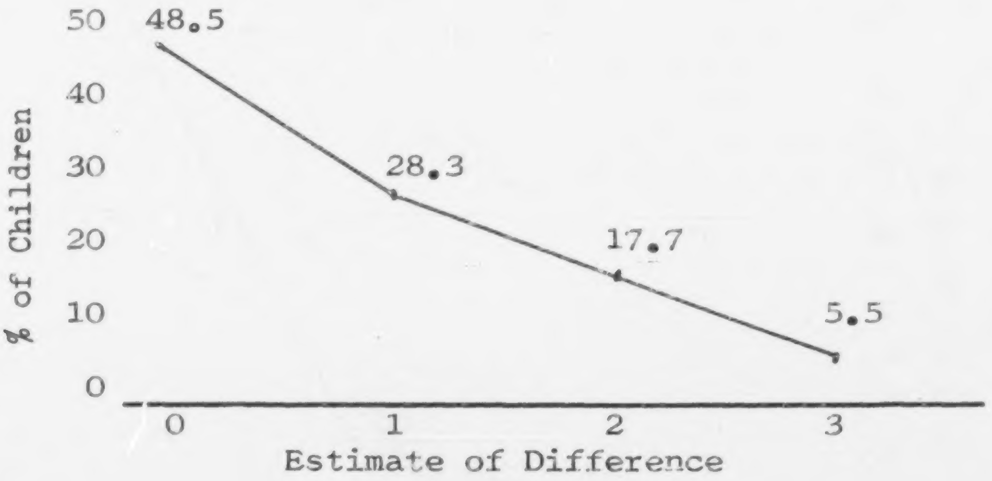


Fig. 6. The % of children who offered differing estimates of a change in volume of 2 decibels



The graphs show that a majority of children agreed in their estimates of zero and 2 db. changes. An aspect of the results which was unusual was that almost half (48.5%) of the children in Fig. 6 believed there was no change when, in fact, there had been a 2 db. change. This is to be compared with the 48.9% in the upper graph who estimated that a small change had occured when there was none. Averaging the percentages of the first two columns of both graphs, we find that 78.1% agreed that the patterns were either unchanged or only slightly changed. It seemed clear that a 2 db. change was insufficient to serve as a cue in identifying the relevant feature.

The results for changes between patterns of 4 and 6 db. are shown in Figures 7 and 8.

Fig. 7. The % of children who offered differing estimates of a change in volume of 4 decibels.

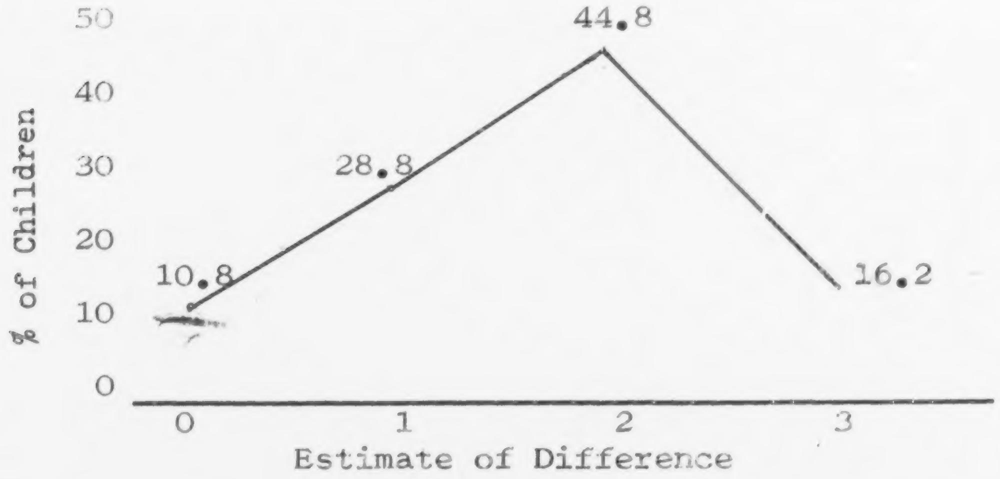
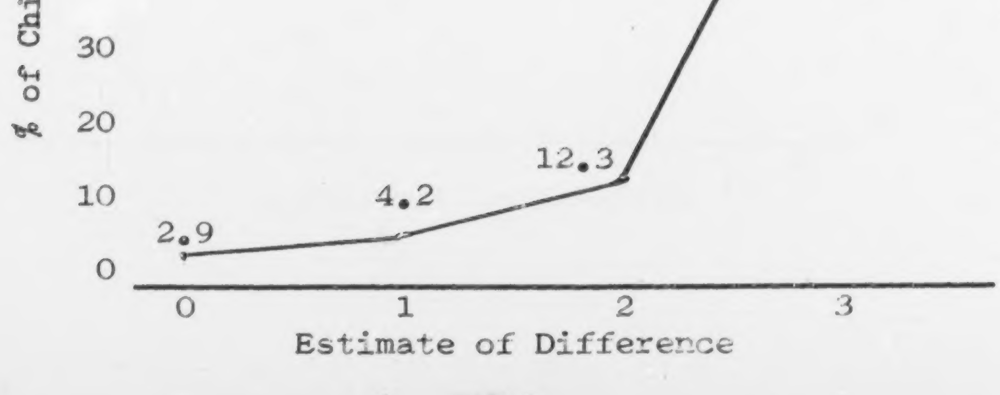


Fig. 8. The % of children who offered differing estimates of a change in volume of 6 decibels.



The high percentage of children who agreed that a 6 db. change was large is remarkable. This was possibly due to the fact that during the training period the extreme example of an increase had been a 6 db. change. Thus, the limit of the range of change had been revealed (Underwood, 1966). These data were used in establishing the values of the relevant and irrelevant features. When volume was a relevant feature, the volume level was held constant for the positive instances and was increased by 6 db. for the negative instance. When volume was irrelevant, there was a 2 db. increment from one instance to the next.

The rather large difference between the values of the relevant feature (6db.) and the irrelevant feature (2db.) was built into the tasks with the intention of making a clear distinction between changes which were relevant and those which were irrelevant. This same procedure was followed with the other features.

Fig. 9. The % of children who offered differing estimates of a change in tempo when there was no change

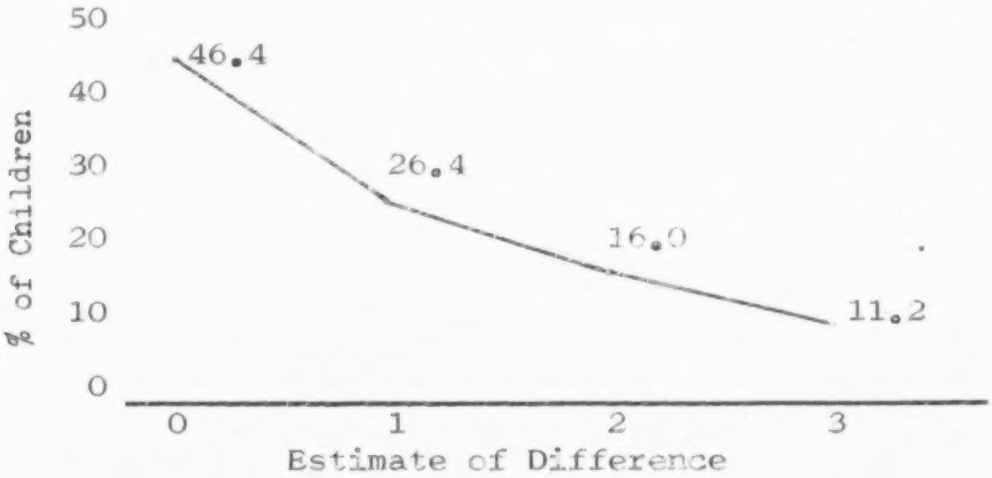
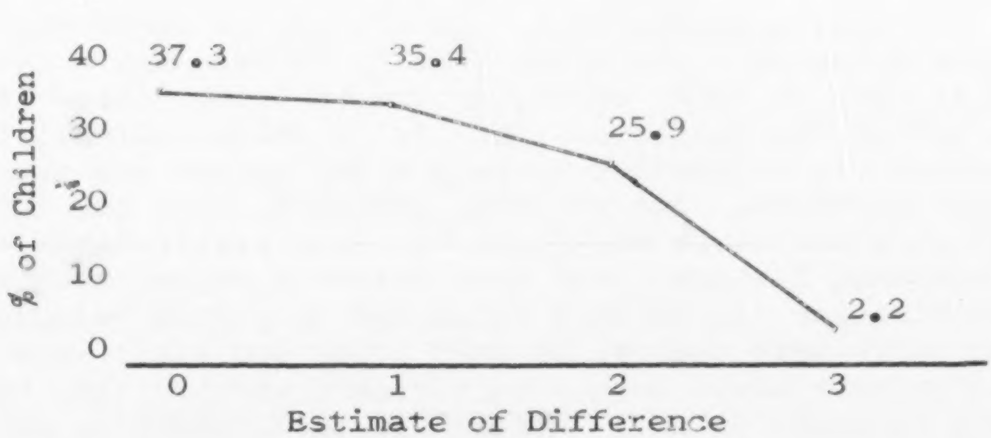


Fig. 10. The % of children who offered differing estimates of the change between 80 and 86 MM



In the case of tempo, each pattern was first played for the children at 80 MM and repeated either without change or by increasing the tempo by 6, 12, or 18 metronome units. Figures 9 to 12 show how children estimated various increments in tempo.

Fig. 11. The % of children who offered differing estimates of the change between 80 and 92 MM

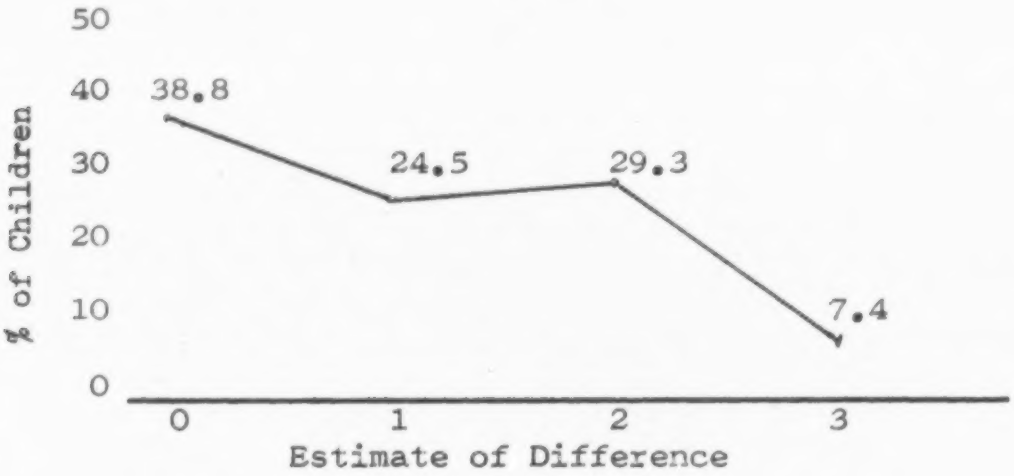
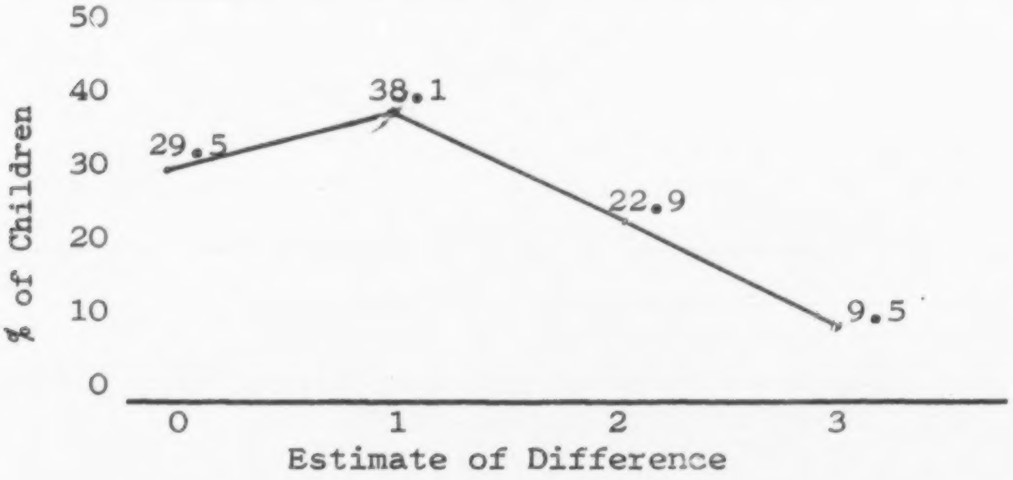


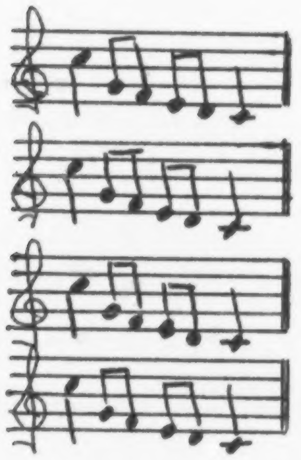
Fig. 12. The % of children who offered differing estimates of the change between 80 and 98 MM



The data indicated that the changes in tempo were not readily detected. One notes that in no case was a tempo change, even as many as 18MM units, estimated to be "large" by more than 10% of the subjects. This is in sharp contrast to the test of volume discrimination where a 6 db. change was regarded as "large" by 80.6%. The evidence suggested that the increments in tempo were not large enough to be used, particularly as cues for the relevant feature. For this reason a second test of tempo discrimination was devised which had 15 paired melodic patterns. Seven pairs were changed by 28MM units and eight were unchanged. Children were asked merely to indicate whether they heard a change in tempo or not. Before the test, examples were played and discussed. The basic finding of this test was that 87.5% of the sample agreed on those melodies which were changed.

Using the data from both tempo tests, the size of tempo increments was set at 28MM units whenever tempo was a relevant feature and 8MM units when it was irrelevant. Two sample tasks are shown in Figures 13 and 14 in which tempo and volume are the variables. In the first task volume is relevant and tempo is irrelevant. The values of the increments are shown at the right of each pattern.

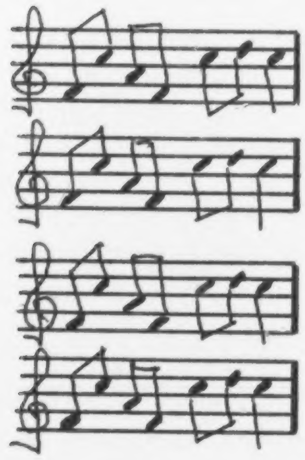
Fig. 13. Volume and tempo changes in sample task



	Relevant Feature Volume	Irrelevant Feature Tempo
1	Base level db	84MM
2	Base level db	92MM
3	+ 6 db	100MM
4	Base level db	108MM

In the second task, tempo is the relevant feature and volume is irrelevant.

Fig. 14. Tempo and volume changes in sample task



	Relevant Feature Tempo	Irrelevant Feature Volume
1	84MM	Base level db
2	112MM	+ 2 db from base
3	84MM	+ 4 db from base
4	84MM	+ 6 db from base

Similar tests of pitch and duration discrimination were developed. In the case of pitch twenty paired patterns were used in which the second pattern was either unchanged, or raised by one, two, or three semi-tones. The findings are shown in Table 2.

TABLE 2

THE PERCENTAGE OF CHILDREN WHO OFFERED DIFFERING ESTIMATES OF THE SIZE OF A PITCH CHANGE WHEN THERE WAS NO CHANGE, ONE SEMI-TONE CHANGE, TWO SEMI-TONES CHANGE, AND THREE SEMI-TONES CHANGE BETWEEN THE TWO PATTERNS

		Subjective Estimates of the Amount of Change			
		No Change	Small Change	Medium Change	Large Change
Objective changes in the two patterns	None	53.6	25.8	15.7	4.9
	One Semi-Tone	36.6	31.1	21.6	10.5
	Two Semi-Tones	22.5	36.9	34.1	16.3
	Three Semi-Tones	18.1	32.6	39.5	9.9

The data show that few children heard any of the changes as being large; however, the last row indicates that 81.9% heard a change of three semi-tones but offered differing estimates of the size of the change. As in the case of tempo, the pitch changes did not appear to be large enough to be acceptable as cues for the relevant variable; therefore, another pitch discrimination test was administered in which the children had to merely determine whether a change occured or not. In this test the range of items was from no change through one, two, three, four and five semi-tones, or from a unison to a perfect fourth. The findings of this test are shown in Table 3. The high percentage of children who recognized changes of two semi-tones or more is worthy of note. Changes of more than two semi-tones are recognized by fewer children but the general level of recognition remained high.

On the basis of these results, the increment between the positive instances and the negative instance was set at three semi-tones. Where pitch was an irrelevant variable, each instance varied by one semi-tone from the other. In other words the melodic pattern was transposed up a minor third when it was the relevant feature, and up a minor second when it was irrelevant.

TABLE 3

PERCENTAGE OF CHILDREN WHO IDENTIFIED CHANGES OF VARIOUS SIZES

Amount of Change	"No Change"	"Change"
None	86.5	13.5
1 Semi-tone	32.5	67.5
2 Semi-tones	10.0	90.0
3 Semi-tones	25.0	75.0
4 Semi-tones	12.5	87.7
5 Semi-tones	20.0	80.0

The quantitative description of the duration of a tone was a greater problem than in the case of the other elements. Ordinary notational values such as ♩, ♪ were found to be too imprecise to be objectively measured. The method used to produce tones of varying length was to use the Moog synthesizer envelope. The envelope controls the beginning, middle, and end of a tone. The concern here was with the end or decay of a tone. By depressing a key and releasing it one could control the amount of decay electronically. If the keyboard touch was held constant, then variation from legato to staccato could be produced simply by regulating the decay of the tone by means of the Moog. Although the duration of the touch could not be quantitatively determined, the decay was. Every effort was made to keep the touch constant for all patterns and vary the decay only. The durational values of the decay were .200 second, .150 second, .100 second, and .075 second.

Using the durational values of the decay portion of the tone only, tasks were constructed in which the duration of the decay varied between the positive and negative instances from .200 to .075 second. When duration was irrelevant the decay portion of the tone was .200, .150, .100, and .075 second across all four instances.

Relationships Among Test Sequences

All tasks within the same test had the same relevant and irrelevant variables. For example, all the tasks in the first test had volume as a relevant variable and tempo as an irrelevant variable. Various combinations of the features resulted in eight different sequences each having a different combination of relevant and irrelevant features. These combinations are listed below:

Test Sequence	Relevant	Irrelevant
1a	Volume	Tempo
1b	Tempo	Volume
2a	Pitch	Duration
2b	Duration	Pitch
3a	Volume	Pitch
3b	Pitch	Volume
4a	Tempo	Duration
4b	Duration	Tempo

With the exception of the relevant and irrelevant features, the tests were the same in other details. They all used the same melodies in the same order. This was done to control for differences in performance which might result if a different set of melodies were used with each test. That melodic patterns may vary in their "goodness" as bearers of conceptual information is a possibility. There is no known way of quantitatively measuring the potential of a melody to elicit conceptual responses. Therefore, if differences between the melodies did exist, they were equally distributed among the tests. Not only was the same set of melodic patterns used in each of the tests, the melodies were presented in the same order each time. This was done so as to equalize whatever influence the peculiar character of the melody might have on performance.

Summary. The Long form instrument consisted of eight tests with each test containing 20 related tasks. The tasks were designed to elicit conceptual behavior as defined by a number of psychologists. Each task consisted of four melodic patterns, three of which were positive instances and one which was negative. A child was asked to find the pattern which was most "different." The size of the variation of relevant and irrelevant features was determined by establishing a subjective scale of discriminability.

Administration of the Instrument

Prior to the administration of the Long form of the instrument principals of the participating schools were oriented to the purposes and procedures of the study. Participating classrooms were identified and teachers were apprised of the nature of the study. Children were then randomly selected and their parents were asked to grant permission for their child to be in the study. The schedule of interviews prepared for each classroom allowed for each child in that class to be available for 20 minute sessions twice a week. The tests were to be completed by each child at the rate of one sequence per session over a period of four weeks. Every school was to provide an interview area which was quiet and immune to disturbance. The area was to be attractive for a child with adequate lighting, ventilation, and furniture. The children were free to move from their classrooms to the area and back without adult supervision.

Before working on the tasks, a child was acquainted with the sound system, purposes of the "listening games," and the "rules of the games." The orientation session usually involved two to seven children to allay anxiety because of the unfamiliar situation and also to economize on time. Sorting of visual materials such as colors and shapes preceeded sorting of auditory materials. The auditory sorting began with the child's name being repeated in various ways. Names were spoken in sets of four with one being louder, faster, higher, or shorter. Following the name sorting was a task similar to the tasks in the sequences. The orientation meeting concluded with the experimenter asking the child if he understood how to play the game and if he would like to play more listening games. All children responded positively.

Interview Procedure. General instructions in the orientation session were always the same; however, in the individual interview the instructions varied depending on the experimental group the child was a member of. Research assistants followed a set of detailed interview guidelines so as to operate uniformly. Constant checks were made to keep interview procedures as unvarying as possible.

Using a cassette, the experimenter played each task once for a child stopping only long enough between tasks to allow the child to respond and to record the response. When a child had five consecutive correct responses, he was excused from further work for that session. Children who failed to reach the criterion were encouraged to keep trying until all 20 tasks had been heard. If a child became frustrated, anxious, or stubborn, he was allowed to stop at a suitable point in the sequence.

A data gathering form was produced for purposes of recording the observations and the responses of the subject. Data sheets were color coded as a visual referent to the experimental group to which the child belonged. For example, pink data sheets were used for the mode 4 (motor response) group. In addition to recording the responses, the experimenter reported which of the four instances had been selected, verbal and motor responses (if necessary), and any other observations of the child's behavior which might have influenced his performance, e.g., illness, fatigue, anxiety, etc. The essential data was later transferred to IBM computer cards.

Feedback. Experimenters answered affimatively when a child correctly identified the negative instance. The child also received a chip. Chips were turned back to the experimenter at the close of a session in return for a prize. Children chose pencils, pads of paper, etc., as their prize. When a child incorrectly identified an instance as "different," he was told "No, but try another game." Experimenters were warned to avoid any gesture, facial expression, or verbal comment which might directly or indirectly influence a child's performance.

Staff. A staff of 14 research assistants was selected for purposes of conducting the interviews. Twelve of the research assistants were college students who were personally known to the director before research began. The other two assistants were women with teaching experience in the lower grades. Research assistants were trained by the director in the use of the sound system, administration of the tasks, and recording of data. The staff was expected to maintain a good relationship with the classroom teacher and others in the schools where they worked. Through daily contact with the research assistants, the director was able to receive feedback concerning the research progress and special problems.

II. DEVELOPMENT OF THE SHORT FORM OF THE INSTRUMENT

Originally the Long form was planned to be used with 240 children between four and six years of age. Shortly after the interviews began, it became apparent that the tasks would be too difficult for four year olds. Faced with these problems, a decision was made to (1) limit the Long form to 126 children in the five, six, and seven year range, (2) develop a shorter form of the instrument, and (3) use the Short form with a new sample of 260 five, six, and seven year old children.

In the administration of the Long form tests research assistants reported that some children seemed to have trouble keeping all four instances in mind long enough to isolate the negative instance. The difficulty was especially evident with the youngest children in the study. That memory is a factor in conceptual behavior had been shown by other studies (Bourne, 1966). Consequently, a new form was developed which reduced the number of instances in each task to two, rather than four. Of course, one instance was positive, the other negative.

In a two-choice task where one instance was "right" and the other "wrong," the sorting problem was made considerably easier. In fact, it might appear that the task was simply a matter of guessing what the experimenter had in mind. On the other hand, if a child was able to identify the correct instance over a series of tasks, he must have discovered a rule which helped him achieve such consistency. Bruner (1956) designed a conceptual task in which subjects "guessed" the correct instance from a display of cards. In Bruner's study subjects were found to use information from each choice as a clue for the following choice. The task sequences of the Short form also required that subjects use information from previous choices in order to achieve accurate choices with consistency. When a certain level of consistency was reached, it seemed proper to infer that the child was classifying the patterns on the basis of a conceptual rule.

Administration of the Short Form. The procedure used in administering the Short form began with a brief training session in which the rule for solving the task was exemplified. First, the experimenter showed the child two pencils, one red and the other blue. He then asked the child to guess which pencil he (the experimenter) liked best. The red choice was positively reinforced. If a child guessed wrong, he was asked to try again. This procedure was repeated with red and white chips, red and black books, etc. When a child was able to pick the correct color consistently, the experimenter asked the child how he knew which object to select. If the child answered correctly, the same procedure would be followed with geometric shapes. Lastly, the experimenter played a pair of melodic patterns. The patterns differed from each other in one feature such as volume. The experimenter told the child which pattern he liked best. In the case of volume, the correct choice was the louder of the two. The patterns were repeated and the child was asked to indicate which of the two the experimenter liked best. This procedure was repeated until the child responded correctly on three successive trials. At this point the child was told that he was going to listen to some other melodies and that he was to guess which of two melodies the experimenter liked best. The auditory portion of the orientation was used each time the child began a different test; however, the relevant feature was changed.

An important difference in the administration of the Short form was that subjects performed on all 20 tasks of each test. There were only four tests with each one embodying a concept of louder, faster, higher, or shorter. The same melodic patterns were used in the Short form as in the Long form.

Another important difference in the design of the Short form was that the irrelevant feature was severely limited in competing with the relevant feature because there were only two instances. The function of the irrelevant feature was considered to be too inconsequential to require the development of sequences which utilized all the possible combinations of relevant and irrelevant features. Following is a list of the sequences which shows the relevant and irrelevant features.

Short Form Test Sequences

Relevant Feature	Irrelevant Feature
Volume	Tempo
Tempo	Duration
Pitch	Duration
Duration	Tempo

The size of change between the positive and negative instances of the relevant feature was the same in both forms. The increment between the two instances on the irrelevant feature was the same for both forms. Correct answers for each task were distributed through the sequence so that no pattern emerged.

Other aspects of administration such as scheduling, recording, data, and feedback were the same with the Short form as with the Long form.

Summary. The Short form consisted of four tests which required the subjects to utilize concepts of louder, faster, higher, and shorter. Each test had 20 related tasks exemplifying a single concept. The task was a two-choice problem in which the subject was to identify the correct pattern on the basis of a feature which had changed. The feature appeared in either the first or second instance. The order of correct choices was irregular so that no pattern emerged. Subjects who demonstrated consistent accuracy in their choices were assumed to have used a conceptual rule such as "the louder one."

III. THE SAMPLE

Characteristics of the Population

The two forms of the instrument were administered to a sample of kindergarten, first, and second grade children from the Lincoln, Seward, and Centennial school districts. In addition to public schools from these districts, St. John, Calvary, and Redeemer Lutheran schools and the University of Nebraska Nursery school took part in the study.

These schools were located in eastern Nebraska communities whose population for this state may be described as large (Lincoln), medium (Seward), and small (Waco, Beaver Crossing, Goehner, Staplehurst). The schools in the study were chosen because of their accessibility and they were believed representative of a broad socio-economic-cultural spectrum. At one end of the spectrum was Pyrtle elementary school (Lincoln) which was in the middle of a new, affluent, residential area, while at the other end was Saratoga elementary school (Lincoln) in an older, racially mixed area. In cultural opportunities children from St. John Lutheran school were possibly most endowed since the school serves as a lab school for Concordia Teachers College. Other children in the sample were not so fortunate in that they came from farms or lower income homes where educational and cultural advantages were quite limited. In general, the sample was believed to fairly represent the population in the schools of the area in terms of socio-economic level, education, intelligence, and cultural level.

Selection and Organization of the Sample

All the schools which were contacted agreed to participate in the study and all parents, except one, gave permission for their child to take part in the study. The sample was a random, stratified sample, i.e., children were randomly selected from within the grade groups of the participating schools so that each grade group was approximately equal in size. Each group was divided into four approximately equal sub-groups. Children were selected for each sub-group at random.

The total number of subjects in the sample was 386. Of this number 126 children were given the Long form of the instrument and 260 were given the Short form. Table 4 shows the divisions of the sample by grade and form as well as the mean ages.

TABLE 4

DIVISION OF THE SAMPLE BY GRADE AND FORM

Grade	Long Form		Short Form	
	Number of Subjects	N's Mean Age	Number of Subjects	N's Mean Age
Kindergarten	42	5.78	86	5.84
First Grade	42	6.89	85	6.84
Second Grade	42	8.04	89	8.00
Total N	126		260	

IV. RESEARCH DESIGN

Children whose ages were five, six, and seven years old were randomly selected from their classrooms to form three equivalently sized groups. Each age group was further divided into four subgroups. Children were randomly assigned to one of the four subgroups. Each sub-group received varying instructions for performing the tasks.

It was hypothesized that the varying instructions would produce differences in performance on the conceptual tasks; hence, the instructions and the particular form of behavior which they were to elicit were designated as instructional modes, or, simply modes. The sub-groups and their modes are shown below:

Group 1 - Mode 1: Discovery

Group 2 - Mode 2: Verbal cue

Group 3 - Mode 3: Verbal response

Group 4 - Mode 4: Motor response

TABLE 5

DIVISION OF SAMPLE BY GRADE, MODE, AND FORM

Grade	Long Form N's				Short Form N's			
	Mode 1	2	3	4	Mode 1	2	3	4
K	9	13	10	10	21	20	23	22
1	10	11	10	11	21	21	20	23
2	12	10	10	10	21	23	22	23

As can be observed, there were twelve cells which could be compared either horizontally or vertically. An analysis of variance was considered appropriate to the design.

Instructions

Children in all groups were given instructions up to the point that the nature of the problem was explained. The subjects in mode 1 cells were given no further instructions but simply asked to identify the pattern which was "different" or which the experimenter "liked best." These children needed to discover for themselves the relevant cue which led to the correct answer.

Subjects in mode 2 cells were given the added instruction to listen for the louder, faster, higher, or shorter sounding melody. Children in mode 3 cells were given no additional instructions but were asked to offer a verbal explanation for their choice. Those in mode 4 cells were asked to tap, sing, hum, or move in response to the melodic patterns. No additional verbal instructions or responses were required. All subjects were given positive reinforcement whenever they responded correctly. The complete instructions as given by the experimenter are found in the appendix.

CHAPTER IV: RESULTS OF THE STUDY

The purpose of this chapter was to present the results of the experiment as they pertain to the three hypotheses of the study. The organization was derived from these hypotheses. The hypotheses being tested were the effect of age on conceptual behavior, the effect of instructional mode, and the effect of the auditory stimulus pattern.

The first part of the chapter brings together the experimental results of the Short and Long forms of the test which were administered to independent, random samples of 260 and 126 children. Since both forms had the same purpose and design, it was believed to be unnecessary to report the data from both forms as long as the data led to the same conclusions. Because the Short form was administered to a larger sample and statistical procedures could be applied to it more readily, the numerical data, unless otherwise indicated, are derived from the Short form. Where there is an incongruity or novel aspect of the results of the Long form, this aspect will be brought out. The Long form results are summarized in the Appendix.

The second part of the chapter covered the interpretation of the results.

I. EXPERIMENTAL RESULTS

Hypothesis One

There is no significant difference between the observed mean scores and a mean which is attained only by chance. The alternate hypothesis was that the obtained means are significantly higher than that which may be expected by chance. Should the alternate hypothesis be accepted, the inference can be made that concept attainment is, to some extent, being measured. The nature of such an inference is probabilistic, and its validity depends on the accuracy of the observations and the consistency of the performance (Brunswik, 1966). Three methods were used to test the hypothesis. First, the binomial model was used to determine the probability of the observed mean. Since each task was a two-choice problem which resulted in either a right or wrong choice, the probability of success on any task by chance alone was 1/2 or .5. The mean of a binomial distribution is the number of trials times the probability of any trial. Since each sequence contained 20 items, the expected mean, if chance alone were operating, is 10. Secondly, a z score for the observed mean was computed using the Sign test. Thirdly, the consistency of correct responses was measured and the probability of any string of correct responses in 20 trials was calculated. The probability was empirically determined using a computer; however, this method produces similar results to the Bernoulli trials formula. Consistency of accurate responses was taken as an important indicator of conceptual behavior. A consistency score, the longest string of correct responses, was determined for each subject; these scores were summed and a mean consistency score was derived.

These statistical procedures provided the basis for estimating the level of conceptual performance by each age group for the four sequences. Table 6 summarizes the data for Test A - Volume Tasks. Column 1 gives the mean number of correct responses by age group; column 2 gives the binomial probability of the observed mean or higher being attained; column 3 shows the z of the Sign test; column 4 indicates the percentage of correct responses; column 5 gives the mean consistency scores; column 6 registers the probability of such a string of correct responses; and column 7 reveals the percentage of children whose consistency score was at or higher than the .01 probability level. Incidentally, this consistency score is nine correct responses in a row. It may be mentioned that it makes no difference where in the 20 trials the nine responses occur, nor does the formula take into account the possibility that the subject may have had other correct responses but with less consistency.

TABLE 6

PROBABILITY OF OBSERVED MEANS ON VOLUME TASKS

	1 Accuracy Mean	2 Binomial Prob.	3 Sign Test z	4 % Correct Responses	5 Consist. Mean	6 Consist. Prob.	7 % at or above p of .01
K	15.3	.02	2.36	76.7	10.0	.006	56.0
1	18.2	.0002	3.66	91.0	15.0	.00006	79.0
2	18.4	.0002	3.75	92.0	15.5	.00006	85.0

All probabilities, except column 2, row 1, were significant beyond the .01 level; therefore, the null hypothesis was rejected. It was concluded that the scores did not arise merely from chance but that "something" was consistently effecting the performance. It was inferred that the "something" was a concept of louder and that the children were able to use their concept at a high level of consistency.

Test B - Tempo Tasks Following the same procedure as with the Volume Tasks, the results of the tempo tasks are presented in Table 7.

TABLE 7

PROBABILITY OF OBSERVED MEANS ON TEMPO TASKS

	1 Accuracy Mean	2 Binomial Prob.	3 Sign Test z	4 % Correct Responses	5 Consist. Mean	6 Consist. Prob.	7 % at or above p of .01
K	13.1	.3	1.38 ns	65.6	6.3	.12 ns	22.0
1	16.6	.005	2.94	83.0	11.3	.003	61.0
2	16.8	.001	3.03	84.0	11.8	.002	67.0

The scores of the kindergarten group were not significantly different than that which could result by chance alone; therefore, the null hypothesis was retained for this level. It should be noted, however, that 22% of the kindergarteners performed at a level which was significantly beyond chance expectations. The first and second grade means were all significant beyond the .01 level; therefore, the null hypothesis was rejected. It was concluded that chance alone could not account for the high scores and that the groups demonstrated the use of a concept of faster.

Test C - Pitch Tasks The results of the pitch tasks are shown in Table 8 together with the probabilities of the observed means.

TABLE 8

PROBABILITY OF OBSERVED MEANS ON PITCH TASKS

	1 Accuracy Mean	2 Binomial Prob.	3 Sign Test z	4 % Correct Responses	5 Consist. Mean	6 Consist. Prob.	7 % at or above p of .01
K	10.8	.59	.35 ns	54.0	4.7	.24 ns	7.0
1	12.6	.25	1.16 ns	63.0	7.0	.05	29.0
2	11.8	.41	.80 ns	59.0	5.8	.2 ns	15.0

With the exception of the first grade consistency mean, the other means were not significantly different from what might be attained by chance alone. Since the first grade consistency mean was significant at only the .05 level, the null hypothesis was retained. It was concluded that conceptual behavior could not be inferred on the basis of these means. However, at each age level there were some children whose consistency scores were significantly different from chance expectations. This was most evident at the first grade level where 29% were at or above the .01 level of probability.

Test D - Duration Tasks Table 9 shows the observed means and probabilities for the durational tasks at each age level.

TABLE 9

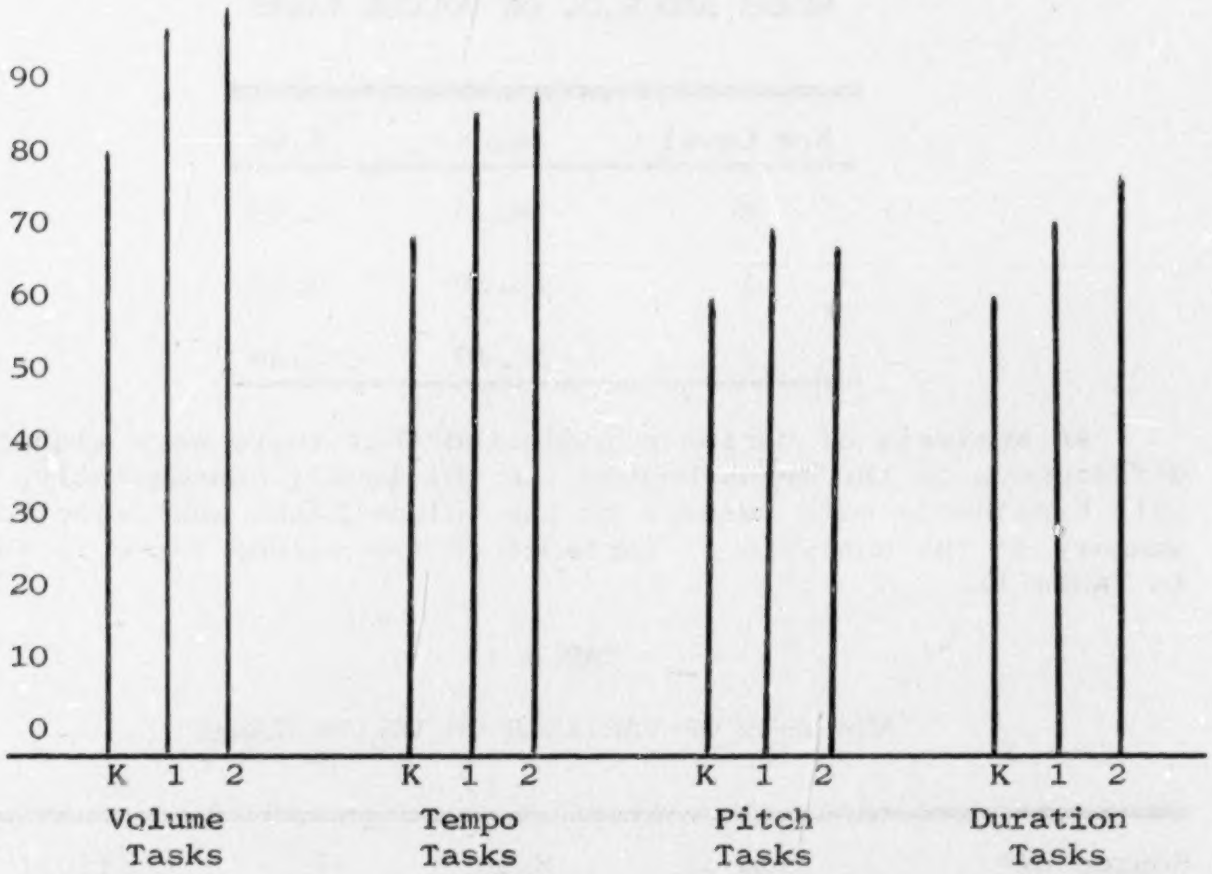
PROBABILITY OF OBSERVED MEANS ON DURATION TASKS

	1 Accuracy Mean	2 Binomial Prob.	3 Sign Test z	4 % Correct Responses	5 Consist. Mean	6 Consist. Prob.	7 % at or above p of .01
K	12.1	.25	.93 ns	60.5	5.7	.10	17.0
1	13.4	.13	1.51 ns	67.0	7.5	.05	32.0
2	14.2	.05	1.83 $p<.05$	71.0	8.1	.02	34.0

The mean scores of the kindergarten group were not significantly different from chance probability; therefore, the null hypothesis was retained. At the first grade level only the mean consistency score differed significantly from chance at the .05 level. Since the mean accuracy score was not significant by either the Sign test or the binomial model, the null hypothesis was retained. However, it was concluded that no strong case could be made in favor of the null hypothesis inasmuch as the two means straddled the fence of significance with one being significantly different and the other not. It should be noted that almost one-third of the first grade group attained consistency scores which were at the .01 level of significance. With respect to the second grade group, both means were significant ($p<.05$), and, consequently, the null hypothesis was rejected. Like the first graders, about one-third of the group had consistency scores which were at or above the .01 significance level. It was concluded that the second grade scores could not be accounted for by mere chance and that a concept of duration was being used. The performance of the kindergarten and first grade groups does not allow a conclusion that consistent conceptual behavior was involved.

A summary of the results on all sequences of the Short form by age levels is shown in Figure 15. Any comparison of the scores of one task with another must be made with reference to the specific test instrument. From these results one can see that pitch scores are the lowest and volume scores are the highest; however, to make the statement that such differences imply general differences of concept attainment is somewhat risky. An important variable

Fig. 15. Summary of percentage of correct responses on each of the sequences by age group



which may affect conceptual behavior is the task design and particularly the size of the difference between the negative and positive instances. An interpretation of the differences in scores will be reserved until later.

Hypothesis Two

The second null hypothesis was that there is no significant difference in the means of scores by different age groups. This hypothesis was tested using the results of volume, tempo, pitch, and duration tests. The statistical procedure which was used to test this hypothesis was a two-way analysis of variance with age as one of the main effects. An F ratio was found for the age and the mode effect as well as the interaction effect. Subsequently, an analysis of the differences between the means of each cell in the sample matrix was made using a procedure recommended by Garrett (1958, p. 281). This procedure identified the specific means which differed significantly from each other.

For Test A - Volume Tasks, the means and standard deviations obtained for each age level are shown in Table 10.

TABLE 10

MEANS AND S.D. ON VOLUME TASKS

Age Level	Mean	S.D.
K	15.35	4.13
1	18.20	2.89
2	18.40	2.44

An analysis of variance indicated that there were significant differences of the means beyond the .01 level; consequently, the null hypothesis with respect to the volume tasks was rejected. A summary of the analysis of variance on the volume tasks is shown in Table 11.

TABLE 11

ANALYSIS OF VARIANCE ON VOLUME TASKS

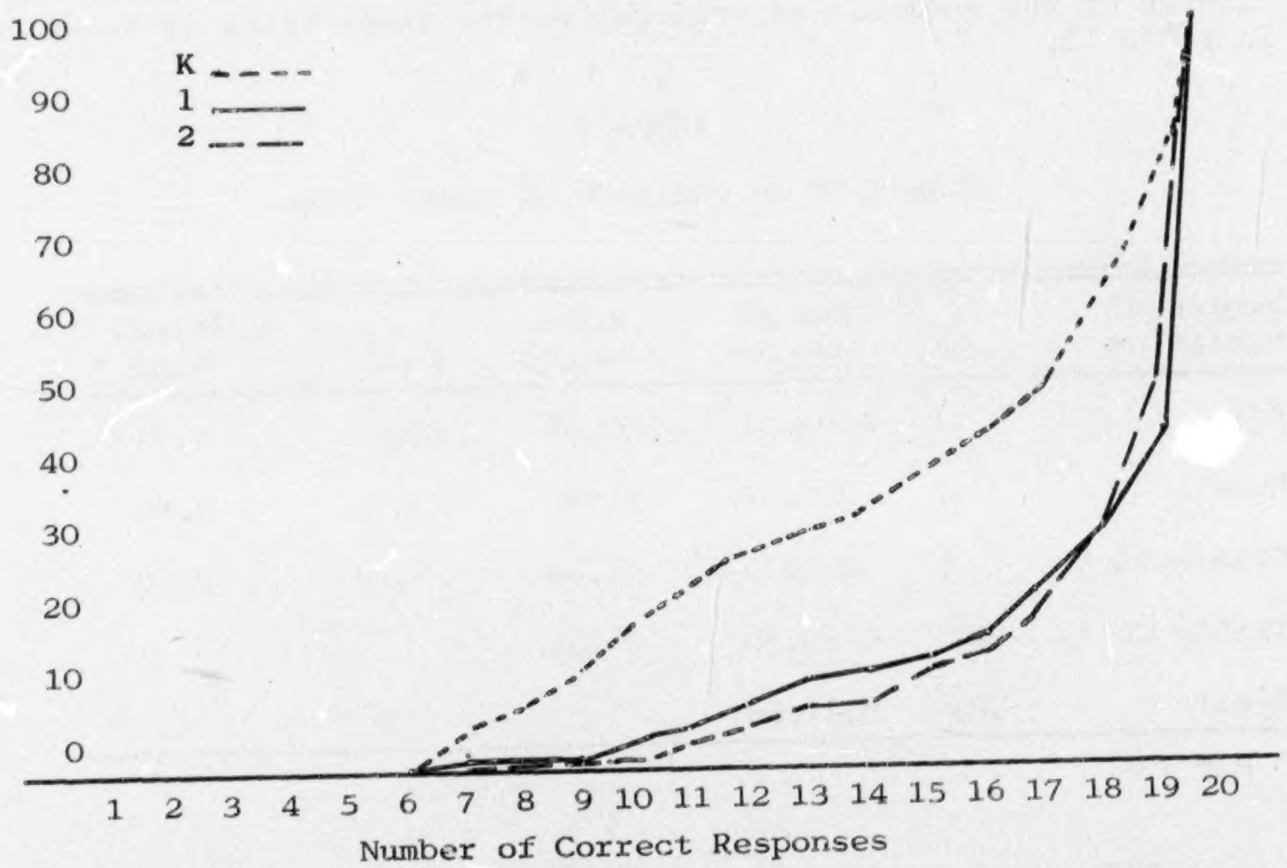
Source of Variation	DF	Sum of Squares	Mean Squares	F Ratio	Critical F Value*
Age	2	504.63	252.31	24.59	4.71
Mode	3	53.52	17.84	1.74	3.88
Interaction	6	78.36	13.06	1.27	2.90
Within Cells	248	2544.68	10.26		
Total	259	3181.21			

* p = .01

It may be noted in passing that the effects of mode and interaction were not significant; however, further discussion of these effects will be reserved until the third hypothesis is considered. The differences in performance between age groups is clearly represented in Figure 16 which shows the percentage of children who achieved a given number of correct responses.

The graph shows that significant differences occured between the kindergarten and first grade levels with no significant differences between the first and second grade levels. The graph also reveals that the distribution of scores is negatively skewed.

Fig. 16. Ogive of the percentage of Ss reaching a given number of correct responses or lower on volume tasks



Most scores occur at the higher end of the range while there are no scores lower than seven. Although there were significant differences between age groups, nevertheless all groups manifested a high performance level. It was concluded that age is a significant factor in these tasks and that the greatest difference was between the kindergarten and first grade levels.

Test B - Tempo Tasks For Test B the means and standard deviations were found for each age level and are shown in Table 12.

TABLE 12

MEANS AND S.D. ON TEMPO TASKS

Age Level	Mean	S.D.
K	13.14	3.93
1	16.63	3.39
2	16.85	3.08

An analysis of variance indicated that there were significant differences between the means beyond the .01 level; consequently the null hypothesis for the tempo tasks was rejected. A summary of the analysis of variance on the tempo tasks is shown in Table 13.

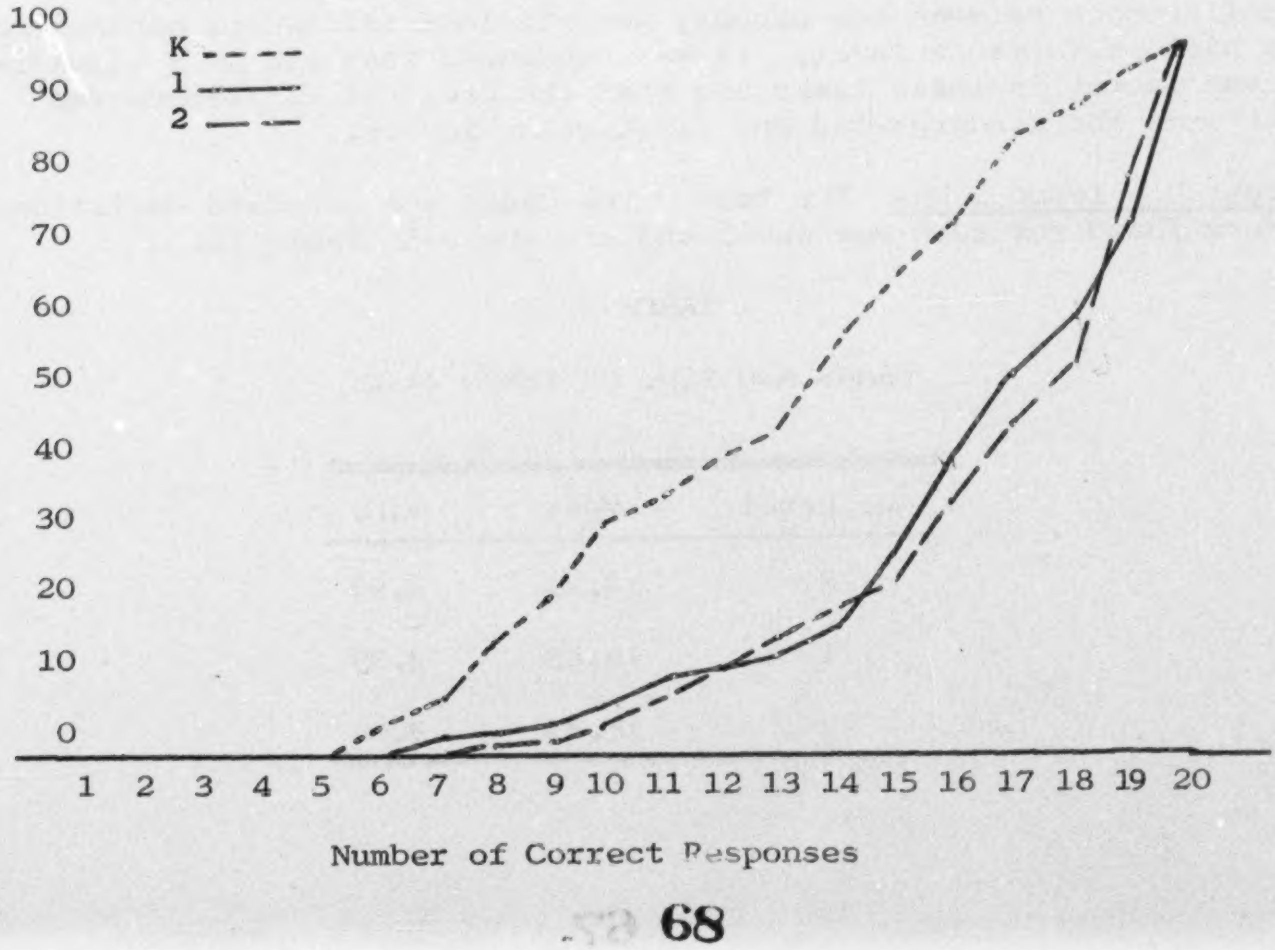
TABLE 13

ANALYSIS OF VARIANCE ON TEMPO TASKS

Source of Variation	DF	Sum of Squares	Mean Squares	F Ratio	Critical F Value *
Age	2	1605.61	802.58	29.66	4.71
Mode	3	275.49	91.83	3.39	3.88
Interaction	6	250.13	41.68	1.54	2.90
Within Cells	248	6710.80	27.05		
Total	259	8841.60			

* p = .01

Fig. 17 Ogive of the percentage of Ss reaching a given number of correct responses or lower on tempo tasks



As can be seen from Table 12, the greatest difference in means was between the kindergarten and first grade levels. This difference was significant at the .05 level. It was concluded that age was a significant factor and that the greatest difference was between the kindergarten and first grade levels. There was very little difference in the means of the first and second grade groups. Figure 17 shows the cumulative percentages of subjects who reached a given score or lower.

Test C - Pitch Tasks For the pitch tasks the means and standard deviations obtained for each age level are shown in Table 14.

TABLE 14

MEANS AND S.D. ON PITCH TASKS

Age Level	Mean	S.D.
K	10.84	3.23
1	12.66	4.06
2	11.77	4.05

An analysis of variance showed that there were significant differences between the means beyond the .01 level; consequently, the null hypothesis for pitch tasks was rejected. A summary of the analysis of variance is shown in Table 15.

TABLE 15

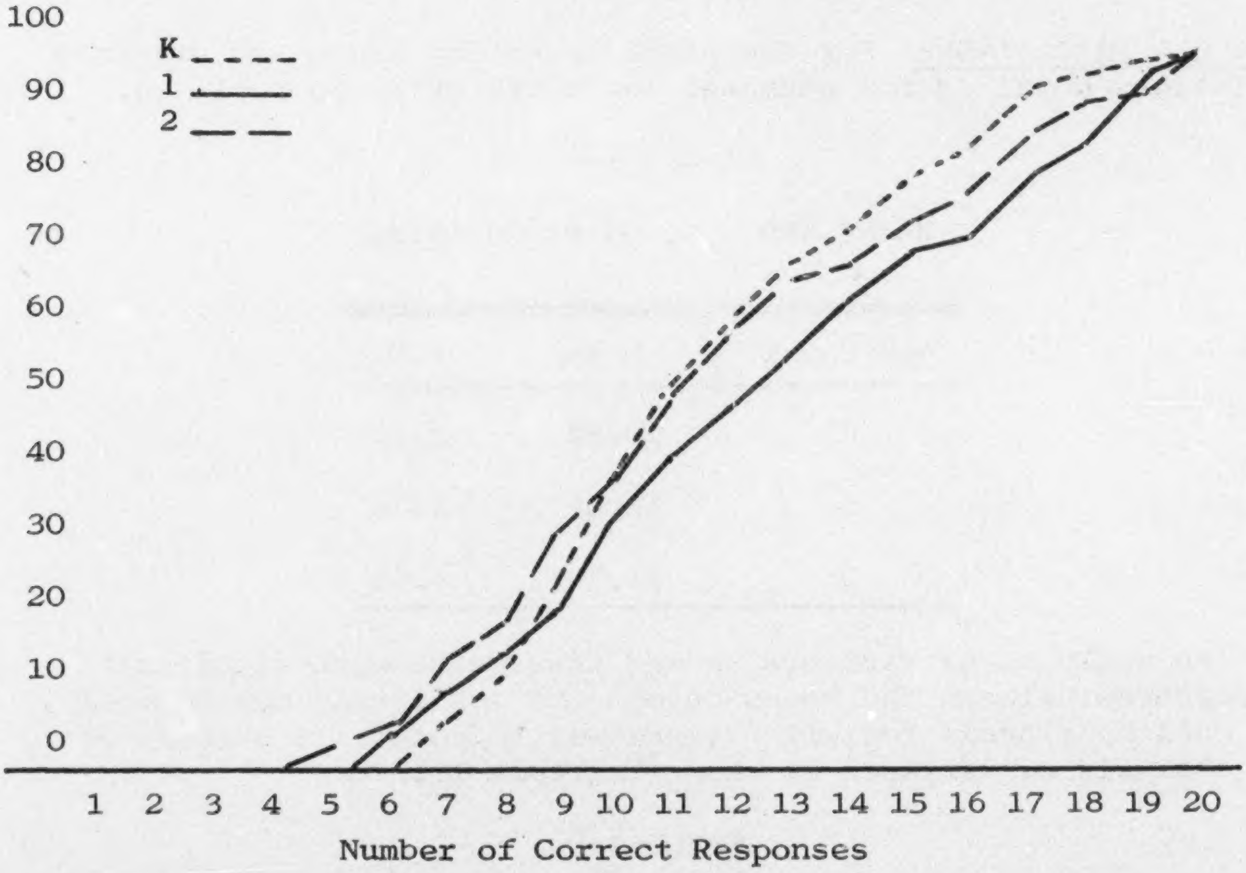
ANALYSIS OF VARIANCE ON TEMPO TASKS

Source of Variation	DF	Sum of Squares	Mean Squares	F Ratio	Critical F Value *
Age	2	208.69	104.34	6.03	4.71
Mode	3	110.54	36.84	2.13	2.13
Interaction	6	242.26	40.37	2.33	2.90
Within Cells	248	4289.71	17.29		
Total	259	4851.21			

* p = .01

It was concluded that age is a significant variable on this test and that the greatest difference ($p < .05$) occured between the kindergarten and first grade levels. It should be noted that the first grade group mean was higher than the second grade mean. This was the only exception to the trend that second grade scores were higher than the other levels.

Fig. 18 Ogive of the percentage of Ss reaching a given number of correct responses or lower on pitch tasks



Test D - Duration Tasks On the duration tasks the means and standard deviations which were obtained for each age level are shown in Table 16.

TABLE 16

MEANS AND S.D. ON DURATION TASKS

Age Level	Mean	S.D.
K	12.08	3.55
1	13.38	4.24
2	14.20	3.88

An analysis of variance showed that there were significant differences between the means beyond the .01 level; consequently, the null hypothesis for duration tasks was rejected. A summary of the analysis of variance is shown in Table 17.

TABLE 17

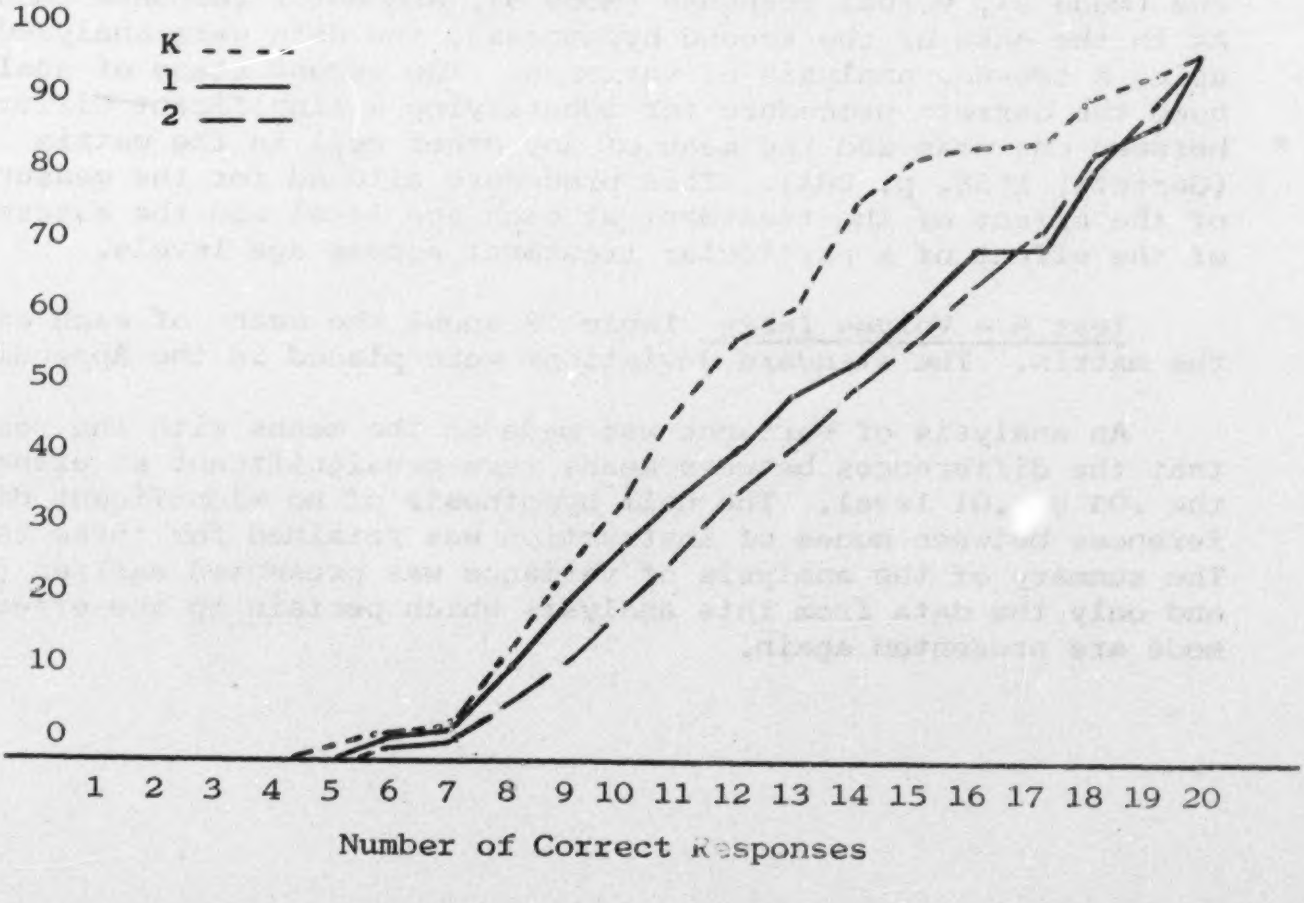
ANALYSIS OF VARIANCE ON DURATION TASKS

Source of Variation	DF	Sum of Squares	Mean Squares	F Ratio	Critical F Value *
Age	2	199.41	99.70	6.519	4.71
Mode	3	89.92	29.97	1.96	3.88
Interaction	6	32.15	5.35	0.35	2.90

* p = .01

Once again the means varied significantly with the greatest variation between the means of the kindergarten and first grade groups. The ogive graph of Figure 19 shows the comparative performance levels of each group.

Fig. 19. Ogive of the percentage of Ss reaching a given number of correct responses or lower on duration tasks



Summary and Conclusions. The results indicated that there were performance differences on each test between the three age levels. These differences were significant in each case beyond the .01 level. The greatest difference of means was between the kindergarten and the first grade levels with the second grade means being slightly higher in all sequences except for the pitch tasks where the first graders did better.

It was concluded that age is a variable in the performance of these tasks and that, generally, the older children score higher than the younger. This conclusion is in agreement with common sense and casual observation of children's performance by their classroom teachers. However, it is not an explanation of these differences in performance since it does not indicate any causal factors. Moreover, it fails to account for individual differences which obviously involve more than age. In short, the fact that differences exist suggests that more particular questions must be raised. As far as this experiment was concerned, the question of the mode of instruction was considered to be a possible variable of conceptual behavior.

Hypothesis Three

The third hypothesis was that there is no significant difference between the means of groups using four different modes of instruction. To test this hypothesis each age level of the sample was divided into four treatment groups. Individuals were placed into each group randomly. The four modes of instruction were discovery (Mode 1), verbal cue (Mode 2), verbal response (Mode 3), and motor response (Mode 4). As in the case of the second hypothesis, the data were analysed using a two-way analysis of variance. The second stage of analysis used the Garrett procedure for identifying a significant difference between one mean and the mean of any other cell in the matrix (Garrett, 1958, p. 281). This procedure allowed for the measurement of the effect of the treatment at each age level and the assessment of the effect of a particular treatment across age levels.

Test A - Volume Tasks Table 18 shows the mean of each cell in the matrix. The standard deviations were placed in the Appendix.

An analysis of variance was made on the means with the result that the differences between means were nonsignificant at either the .05 or .01 level. The null hypothesis of no significant differences between modes of instruction was retained for these tasks. The summary of the analysis of variance was presented earlier (p. 60) and only the data from this analysis which pertain to the effect of mode are presented again.

TABLE 18

MEANS FOR EACH MODE AT EACH AGE LEVEL - VOLUME TASKS

Age	Mode 1	Mode 2	Mode 3	Mode 4
K	15.33	17.10	15.39	13.73
1	17.90	18.29	17.95	18.61
2	18.14	18.78	18.45	18.22
Mean of Means by Mode	17.13	18.09	17.22	16.90

TABLE 19

ANALYSIS OF VARIANCE OF COLUMN MEANS - VOLUME TASKS

Source of Variation	DF	Sum of Squares	Mean Squares	F Ratio	Critical F Value *
Mode	3	53.52	17.84	1.74	2.65

* p = .05

Further comparison of the means of the individual cells with one another was made to determine the possible effect of a mode within a particular age level. The only significant difference was at the kindergarten level where the mean of Mode 2 - Verbal Cue differed significantly ($p < .01$) from the mean of the Motor Response - Mode 4. Analysis of the means of the first and second grade levels showed no significant differences between the means of the cells at either level. It was concluded that the instructional mode did not make any significant difference in the performance of the first and second grade groups.

An examination of the means of the columns shows that, with one exception, there was a large difference between the kindergarten and first grade means. These differences are all significant at the .01 level. The exception is the Verbal Cue - Mode 2 column. In this column the means of the kindergarten cell (17.10), the first grade cell (18.29), and the second grade cell (18.78) did not differ significantly from each other. In other words, the verbal cue positively influenced the performance of the kindergarten subjects in the Mode 2 cell so that the observed mean did not differ significantly from the others. This was not the case for the other modes.

It should be pointed out that, in general, the means of Mode 2 and Mode 3 cells are somewhat higher than the other modes. While these differences are not always significant, there is a definite trend toward higher scores in the cells of Modes 2 and 3 which can be noted in other tests also.

A further analysis was made of the verbal reports of the children in Mode 3. Table 20 summarizes the verbal reports given by children at the different age levels.

TABLE 20

VERBAL REPORTS ON VOLUME TASKS

Verbal Response	Kindergarten		First Grade		Second Grade	
	Freq.	%	Freq.	%	Freq.	%
"Louder" (Over 75% of the time)	15	65.2	15	75.0	20	90.9
"Louder" (25% to 75% of the time)	2	8.7	1	5.0	2	9.1
Total Correct	17	73.9	16	80.0	22	100.0
Incorrect Responses	6	26.1	4	20.0	0	0

The interesting aspect of the verbal response analysis is that the difference in the correct verbal responses between the kindergarten and first grade level is only 6%, whereas the first grade differs from the second by 20%. This is just the reverse of the scores obtained for the test itself where the first grade did not significantly differ from the second. It would appear that the first grade group was able to perform nearly as well on the nonverbal tasks but, when asked to justify their choice verbally, were unable to do so much better than kindergarteners.

Test B - Tempo Tasks Table 21 shows the means of each cell in the matrix obtained for Test B.

An analysis of variance of the column means showed that there was a significant difference between the means beyond the .01 level. Therefore, the null hypothesis of no difference was rejected. An abbreviated analysis of variance as it pertains to the column means is shown in Table 22.

TABLE 21

MEANS FOR EACH MODE AT EACH AGE LEVEL - TEMPO TASKS

Age	Mode 1	Mode 2	Mode 3	Mode 4
K	12.33	14.75	13.30	12.27
1	14.81	16.95	17.85	16.96
2	16.19	18.30	17.00	15.87
Mean of Means by Mode	14.44	16.75	15.95	15.07

TABLE 22

ANALYSIS OF VARIANCE OF COLUMN MEANS - TEMPO TASKS

Source of Variation	DF	Sum of Squares	Mean Squares	F Ratio	Critical F Value *
Mode	3	195.19	65.06	5.67	3.88

* p = .01

Having established that there were significant differences between the means of the various modes, the next step was to determine more precisely where the significant differences were. An analysis of the cells of the kindergarten level showed that the mean for Mode 2 - Verbal Cue was significantly higher (.05 level) than the means of Mode 1 - Discovery and Mode 4 - Motor. There was no significant difference between the two verbal modes. At the first grade level the Mode 1 mean was significantly lower than the means of the other cells ($p < .05$); however, the other cells were not significantly different from each other. At the second grade level the Mode 2 - Verbal Cue mean was significantly higher ($p < .05$) than the means of Modes 1 and 4. There was no significant difference between Modes 2 and 3. In sum, there was a significant difference between the instructional modes with the verbal cue mode being higher than the discovery mode at all three levels and higher than the motor mode at the kindergarten and second grade levels.

The verbal report of the children in the Mode 3 cells is shown in Table 23.

TABLE 23

VERBAL REPORTS ON TEMPO TASKS

Verbal Response	Kindergarten		First Grade		Second Grade	
	Freq.	%	Freq.	%	Freq.	%
"Faster" (Over 50% of the time)	13	56.5	18	90.0	18	81.8
"Faster" (Under 50% of the time)	2	8.6	0	0	1	4.5
Total Correct	13	65.1	18	90.0	19	86.3
Incorrect Responses	8	34.8	2	10.0	3	13.5

The accuracy of the verbal responses closely matches the means obtained on the non-verbal tasks in which kindergarten level was the lowest and the first and second grades were close together. It was apparent that the older children were able to verbalize their concept of faster at a high level of consistency.

Test C - Pitch Tasks Table 24 shows the means of each cell in the matrix for the pitch tasks.

TABLE 24

MEANS FOR EACH MODE AT EACH AGE LEVEL - PITCH TASKS

Age	Mode			
	1	2	3	4
K	10.38	11.10	11.09	10.77
1	12.29	11.38	13.55	13.39
2	9.90	13.13	13.36	10.61
Mean of Means by Mode	10.86	11.92	12.62	11.60

An analysis of variance on the column means revealed that there was no significant difference between the means; consequently, the null hypothesis was retained. An abbreviated analysis of variance of the column means is shown in Table 25. See Table 15 for complete summary of the analysis.

TABLE 25

ANALYSIS OF VARIANCE OF COLUMN MEANS - PITCH TASKS

Source of Variation	DF	Sum of Squares	Mean Squares	F Ratio	Critical F Value *
Mode	3	102.25	34.08	2.45	2.65

* p = .05

Further analysis of the means at each age level showed that there was no significant difference between the obtained means at the kindergarten and first grade levels. There was, however, a significant difference between the means at the second grade level. The verbal modes (Modes 2 and 3) differed significantly (.05 level) from the discovery and motor modes. Generally, the mode of instruction did not appear to affect performance on this sequence significantly with the exception of the second grade level. The interpretation of such significance must be made in the light of the rather low means of Mode 1. This mean was lower than the kindergarten mean of Mode 1; therefore, the significance is not that Mode 2 mean was so high but that the Mode 1 mean of the second grade level was so low.

A comparison of the means of the pitch tasks of the Short form with pitch tasks of the Long form shows an interesting reversal of a trend. With few exceptions the verbal cue mode cells have had higher means than the non-verbal cells. In the two sequences of the Long form which diagnosed performance on pitch tasks, the verbal cue cells were lower in each case. Although the lower level of performance was not significantly different from the non-verbal modes, the fact that there was a reversal of a trend on pitch tasks bears further investigation.

Turning to the verbal reports given by children in the Mode 3 cells, we find that comparatively few children were able to describe the difference in pitch using the words "higher" or "lower." Table 26 summarizes the verbal reports at the three age levels.

The following points should be noted. First, there was little difference in the accuracy of the verbal response and the non-verbal tasks at the first and second grade levels; however, the first graders did better in the non-verbal tasks whereas the second graders did better in giving an accurate verbal report. Secondly, all age levels gave more incorrect verbal responses than correct ones. This same relationship held in the case of the Long form also. On both forms the highest percentage of children used "louder" to designate the change which they heard. In the case of the Long form

TABLE 26

VERBAL REPORTS ON PITCH TASKS

Verbal Response	Kindergarten		First Grade		Second Grade	
	Freq.	%	Freq.	%	Freq.	%
"Higher" (Over 50% of the time)	2	8.6	7	35.0	10	45.4
"Higher" (Under 50% of the time)	0	0	1	5.0	0	0
Total Correct	2	8.6	8	40.0	10	45.4
Other Responses						
"Louder"	6	25.9	3	15.0	2	9.0
"Faster"	1	4.3	2	10.0	1	4.5
"Don't know"	14	60.7	7	35.0	9	40.8
Total Incorrect	21	89.9	12	60.0	12	54.3

between 50% and 60% of the kindergarten group responded with "louder" to the pitch change; no kindergarteners used the label "higher."

Test D - Duration Tasks The means for each cell in the matrix are shown in Table 27.

TABLE 27

MEANS FOR EACH MODE AT EACH AGE LEVEL - DURATION TASKS

Age	Mode			
	1	2	3	4
K	11.76	12.55	12.04	12.00
1	13.81	14.10	13.25	12.43
2	14.10	15.74	13.32	13.61
Mean of Means by Mode	13.22	14.20	12.85	12.69

An analysis of variance on the column means showed that there was no significant difference between the means; therefore, the null hypothesis was retained. The abbreviated analysis of variance of the column means is shown in Table 28. See Table 17 for complete summary.

TABLE 28

ANALYSIS OF VARIANCE OF COLUMN MEANS - DURATION TASKS

Source of Variation	DF	Sum of Squares	Mean Squares	F Ratio	Critical F Value
Mode	3	89.92	29.97	1.96	2.56 (.05) 3.88 (.01)

Further analysis of the cells in each row revealed no significant differences between the modes of instruction for the kindergarten and first grade levels. At the second grade level, the verbal cue mode was significantly higher ($p < .05$) than the verbal response cell.

Analysis of the verbal responses showed that only three of the second grade children used the word "shorter" to describe the staccato sound. Other related terms were "jumpy," "jerky," "bumpy," or "space between the notes." A summary of these responses is shown in Table 29.

TABLE 29

VERBAL REPORTS ON DURATION TASKS

Verbal Response	Kindergarten		First Grade		Second Grade	
	Freq.	%	Freq.	%	Freq.	%
"Shorter"	0	0	0	0	3	13.6
Synonym for Shorter	12	52.2	9	45.0	10	45.4
Total Correct	12	52.2	9	45.0	13	60.0
Other Responses						
"Louder"	2	8.6	1	5.0	3	13.6
"Faster"	1	4.6	4	20.0	2	9.0
"Higher"	0	0	1	5.0	1	4.3
"Don't know"	8	34.7	5	25.0	3	13.6

There did not seem to be any significant difference between the different age levels in the accuracy of applying a verbal label. None of the groups showed a level of consistency above 60%.

Summary and Conclusions Pertaining to the Third Hypothesis

The null hypothesis was sustained for each sequence except for Test B - Tempo Tasks. The mode of instruction appeared to be less of a factor than the age level. In some cases the mode was effective at one age level but not another level, or it was more effective with one kind of conceptual problem than another. Generally, the verbal cue resulted in higher means than the other modes. This was particularly true for the tempo tasks at each of the age levels; it was also true for volume tasks at the kindergarten level and for pitch and duration tasks at the second grade level. The exception was noted with reference to the Long form pitch tasks where the verbal cue resulted in lower means than the non-verbal modes. The surprising thing about the means of the verbal cue mode was not that they were higher (this is what common sense might expect): the surprise was that the verbal cue did not make a significant difference more often. It should also be noted that children tended to perform at a higher level on the non-verbal tasks than in providing accurate verbal reports of the tonal changes. This was particularly true for the kindergarten level while the distinction between a pre-verbal concept and a verbal concept became finer at the first and second grade levels.

Reliability

The reliability of the test measures was estimated using the split-half technique (odd-even) and applying the Spearman - Brown prophecy formula. A second estimate of reliability was obtained from the Kuder - Richardson formula 21. Both methods of estimating reliability were considered appropriate for this test battery since each sequence was designed as a homogenous test. Guilford (1950) has pointed out that the Kuder - Richardson and the Spearman - Brown formulas call for items of equal, or nearly equal, difficulty and intercorrelation. An item analysis of the results showed no wide departure by any item from the average of all items in the sequence. (The item analysis is found in the Appendix.)

The results of these estimates of reliability are shown in Table 30.

Since no results were available from comparable measures, these estimates of reliability must be interpreted as they stand. The two formulas yielded estimates which were generally similar to each other. It was concluded that the reliability of the test measures was within an acceptable range.

TABLE 30

RELIABILITY OF THE TASK TESTS

Task-Sequence	Age	Split-Half	K-R 21
Volume	K	.88	.83
	1	.90	.85
	2	.84	.80
Tempo	K	.71	.75
	1	.80	.80
	2	.83	.76
Pitch	K	.51	.55
	1	.82	.76
	2	.70	.75
Duration	K	.71	.65
	1	.84	.80
	2	.80	.77

Validity

The validity of the test measures was based both on construct validity and external validity. The various constructs and indicators of conceptual behavior have been described in Chapter Two and the design of the tasks incorporating these constructs has been detailed in Chapter Three. The probability of chance alone accounting for the level of performance has been discussed in the present chapter.

To externally validate the measures required some external measure which could be compared with the test results. Unfortunately, there was no standardized test available for this purpose. Consequently, the subjective evaluation of the classroom teacher was used as the basis of external validity. Teachers were asked to rate each child's musical ability using a five point scale. Each teacher was provided with the following description of musical ability and asked to judge the child in terms of this description.

> Musical ability: The ability to move to the beat in music, to indicate pitch changes with the hand, to sing or play melodies with reasonable accuracy, to respond to the emotional content of music.

The estimate of validity was made by correlating the teacher ratings with the test results using the Peason Product - Moment formula. The correlation coefficients are found in Table 31. The numbers in parentheses indicate the correlation between the consistency score and the teacher rating. The other numbers show the correlation of the accuracy score and the teacher rating.

TABLE 31

CORRELATION OF TEACHER RATINGS AND TEST RESULTS

Age	Sequence			
	Volume	Tempo	Pitch	Duration
K	.16 (.11)	.35 (.35)	.28 (.30)	.36 (.29)
1	.24 (.28)	.29 (.30)	.19 (.25)	.10 (.08)
2	-.06 (-.02)	.35 (.34)	.42 (.40)	.27 (.17)
Mean	.11 (.12)	.33 (.33)	.27 (.32)	.24 (.16)

For the size of the sample an r of .25 was considered significant. The highest correlation was found to be .42 on the pitch tasks at the second grade level; the lowest was -.06 on the volume tasks at the second grade level. Over all, the highest correlations were on the tempo tasks and the lowest correlations were for volume tasks.

The limitations of the teacher ratings were fairly obvious. In the first place, these evaluations tended to be global, whereas the tests are specific; secondly, the evaluations were impressionistic; thirdly, they were biased; fourthly, the rating scale covered too small a range and did not allow for sufficient discrimination. A better teacher evaluation procedure could have been devised which would have included a larger range and a continuous scale. The chief advantage of the evaluation procedure was its simplicity. It was clear that teachers tended to over-rate the children from an analysis of the distribution of the ratings which showed a definite skewedness towards the upper end of the scale. Very few children were rated at the lowest end of the scale. In fact, there were no low ratings for the first grade group and only one low rating for the second grade group. Obviously, the teacher ratings themselves are open to question; therefore, the correlation coefficients obtained must be understood as inconclusive.

Other Results

The correlation between the accuracy scores and the consistency scores was determined. The obtained correlation coefficients for each test were as follows: Volume .86, Tempo .84, Pitch .82, and Duration .82.

A correlation between the accuracy scores of each test was made and is shown in Table 32.

TABLE 32

TEST CORRELATIONS

Volume Test	with Tempo Test	.155
	with Pitch Test	.033
	with Duration Test	-.058
Tempo Test	with Pitch Test	.282
	with Duration Test	.123
Pitch Test	with Duration Test	.261

It should be noted that the highest correlations were between the tempo test and pitch test, and pitch test and duration test. The lowest correlations occured between volume and other tests. Duration correlated negatively with volume. None of these correlations were so high as to indicate a close relationship between the two behaviors.

II. INTERPRETATION OF THE RESULTS

The Effect of the Stimulus Variable

It has been shown that there was considerable discrepancy between the performance levels achieved on different kinds of conceptual tasks. At the upper end of the scale were volume tasks; more in the middle were tempo and duration tasks; and at the lower end were pitch tasks. Why should there be such wide disparity in these results? Of particular concern was the difference between the pitch and volume tasks.

There are a number of possible explanations for this large performance discrepancy. One is that children have more experience with volume than with pitch phenomena. If one considers the world of children's experience beyond the sphere of music, it is probable that loud and soft sounds in play, in the home, and elsewhere call for adaptive behavior more often than high or low sounds. Consider, for example, the necessity of speaking loudly enough to be heard in everyday communication. Volume levels of loud and soft tend to be socially reinforced, whereas pitch levels are seldom important in our culture in normal communication. Also, a loud sound is usually a near sound and may stimulate fear, defense, flight, and so on by the child. Warning sounds are usually loud and near sounds; pitch, as a separate entity, has little to do with warning. No child is afraid of a high-pitched bark at a distance, but a loud bark is something else. Then, too, loud sounds generate an affective response which seems to carry over into adult life. Mursell has mentioned that volume was the most primitive tonal phenomenon to which the child responds (Mursell, 1937).

Another explanation is that sound is initially perceived as a whole and only later analyzed according to its components. In many cases volume changes consistently with pitch changes. Ascending pitches are frequently performed at higher levels of intensity and, conversely, the lower sound is often a softer sound. In everyday life the excited child or adult speaks louder and higher, whereas the voice of a calm, relaxed speaker is both softer and lower in pitch. Sachs (1965) has discovered this same tendency in primitive music.

Psychophysical studies of pitch perception by Stevens have shown that pitch discrimination is partly a function of the volume of the stimulus. Within certain pitch ranges a louder sound is heard as a lower sound even though there is no change in the frequency (Stevens, 1938).

A third explanation is based on the inconsistent or incorrect use of verbal levels for pitch and volume stimuli. In the first place, "loud" is almost always a term which refers to auditory stimuli, although, in rare cases, it may be used with reference to color. The terms for pitch, such as "high," "up," "down," and "low," usually refer to a spatial relationship in the environment and become associated with spatial concepts to a much greater extent than to pitch concepts.

Furthermore, it is not unusual in ordinary experience in the home or even at school to refer to loud sounds as high sounds or to soft sounds as low sounds. For example, a child may be told to "lower his voice" or to "turn up the TV."

Most likely the three explanations highlight factors which function together in developing concepts of volume or tend to inhibit concepts of pitch. It is not surprising that the volume concepts tend to dominate and that volume is a more salient attribute. Bearing in mind the tendency for loud and high to have a similar meaning for a child, a teacher need not be dismayed when she instructs the uncertain singer to sing "higher" only to have him sing out-of-tune more loudly.

The Effect of the Age Variable

Earlier it was concluded that age had a significant effect on the performance of these tasks. This conclusion was in agreement with other research studies by Pflederer (1970) and Petzold (1969). Further delineation of the results showed that the kindergarten group differed significantly from the first grade group on each task sequence. Since the mean age of these two groups was 5.8 and 6.8 years, one can say that the greatest change in performance occured between the age of six and seven years. Although the absolute level of performance varied with each task sequence, the gap between the kindergarten and first grade scores prevailed throughout. This constant relationship raised the question of why first and second grade groups varied only slightly from each other while the kindergarten scores were significantly lower. The issue is whether we are dealing with a behavior pattern which is particular with these tasks or whether this behavior is but one facet of a more fundamental difference in children's thinking.

One answer to the question comes from Piagetian theory. Piaget believes that the child passes from the pre-operational period to the concrete operational period about age seven. The difference between the two periods is a qualitative difference in thinking. The concrete-operational child

> seems to have at his command a coherent and integrated cognitive system with which he organizes and manipulates the world around him. Much more than his younger counterpart, he gives the decided impression of possessing a solid cognitive bedrock, something flexible and plastic and yet consistent and enduring, with which he can structure the present in terms of the past without undue strain and dislocation, that is, with the ever-present tendency to tumble into the perplexity and contradiction which mark the preschooler.
> (Flavell, 1963, p. 165)

Piaget stresses that operations cover a broad range of related cognitive acts, not merely logical or numerical operation. One kind of operation is responding to a set of things as similar from some point of view, thereby constituting a class (Flavell, 1963). The child who consistently identifies the higher of two melodic patterns over a series of tasks is forming a class and is demonstrating operational thought. From this premise one may argue that the younger, pre-operational child has not reached the stage of development where he can operate consistently and logically on the different conceptual tasks. He tends to focus on irrelevant aspects and be unable to differentiate between the relevant and irrelevant attributes. This was strikingly demonstrated by the tendency of kindergarteners who centered on volume changes when they were irrelevant or not even present in the task situation. Piaget considers the change from one stage to another a result of interaction between general experience and maturation. What must be understood is that the different stages of intellectual development are changes in the form of thought. Piaget is not particularly concerned with detailing the contents of the child's mind and admits that these contents vary from child to child depending on the particular experiences he has had. What is suggested is that the pre-operational child thinks in a way which is qualitatively different from the operational child. This theory may help to explain the difference in performance between the three age levels.

An alternate theory is that the younger children had fewer musical experiences and thus had less developed concepts. This may be true, but it fails to explain why there was such a slight difference between the first and second grade performance. If musical experience were the only factor, one should observe a smooth improvement from one age to the next. One attractive aspect to the theory is that it explains, in a simplistic way, individual differences. It has been shown that at each level some children performed consistently enough to demonstrate conceptual functioning. Experience is usually assumed to account for these individual differences but as yet we know too little about the kinds of experiences which have contributed to this conceptual development.

Regardless of the theory, one should not conclude that there is no way in which concept development may be further developed at the kindergarten or other age levels. These results by no means specify the fixed boundaries of conceptual development for these ages. Further research is needed whereby the facilitating aspects of instructional programs can be measured in terms of conceptual development.

The Effect of the Mode Variable

Although the effects of mode of instruction were not as differentiated as those of age and kind of task, nevertheless, they may imply certain instructional strategies. The most important aspect of the results seemed to be the relationship between the verbal and non-verbal modes of instruction. In the first place, the verbal modes, both verbal cue and verbal response, were generally higher, and in several cases significantly so, than the non-verbal modes. Secondly, in eleven out of twelve cases there was no significant difference between the verbal cue and verbal response mode. It can be said that the children who were told exactly what attribute to listen for generally did not vary significantly from those who were asked to provide the experimenter with their own explanation of the change they heard.

The fact that the verbal cue mode was effective on certain tasks or at certain age levels is at first glance somewhat mystifying. For example, the verbal cue was most effective on tempo tasks and of little effect on volume and durational tasks. In the case of the Long form, the verbal cue means were sometimes lower than the non-verbal means. Furthermore, the verbal cue was usually more effective with the older, second grade children, than with the kindergarten children. The verbal cue, however, did not appear to benefit the first and second graders when they achieved their highest scores on the volume tasks. On the other hand, the verbal cue did not help kindergarteners on those tasks where they scored rather low, i.e., pitch and duration.

One explanation of these results is that the verbal cue becomes effective only at a particular stage of concept development. At the earliest stage of concept development when the concept is still fragile and labile, a verbal cue may have little or no meaning because it has not become associated with the concept. In some cases the word may even have a detrimental effect on the solving of the conceptual problem. Witness the results of the Short and Long forms pitch tasks. The results of the Short form showed that the verbal cue did not make a significant difference in performance of the kindergarten children; the results of the Long form which combined the irrelevant features of volume and duration with pitch showed that the verbal cue mode performance was lower, sometimes significantly so, than the non-verbal modes. It may be inferred that the instruction "Listen for the higher sounding melody" led some of the children to focus on the louder, irrelevant feature, instead of the pitch difference. In sum, a cue word which has little meaning for the child may not result in a performance which is significantly different from a non-verbal approach, but a word which is misunderstood may have a counter-productive effect on the child's ability to solve the problem.

On the other hand, when a concept is well-established and has reached the verbal concept stage, an external verbal cue may be unnecessary. This does not mean that words cease to exercise any power at this stage. What is more plausible is that the subject has so internalized the relationship between the word and his concept that he provides his own verbal stimulus and derives no further benefit from an outside cue. The closeness of the scores between the verbal cue and the verbal response modes with a well-established concept such as volume would seem to support the above statement.

At some point in concept development the verbal cue seems to be helpful. It helped kindergarteners in volume tasks, it helped all age groups in tempo tasks, and it helped second graders in pitch and duration tasks. Thus, it may be that before a verbal cue can be of much help, a child has to have attained a fairly stable concept. There may be a critical point or phase of conceptual development where a verbal cue is most beneficial in solving conceptual tasks; before this point is reached, the verbal cue is meaningless or misunderstood, after it has passed the verbal cue is redundant.

CHAPTER V: SUMMARY AND CONCLUSIONS

The purpose of this study was to investigate the development of certain concepts of music in kindergarten, first, and second grade children.

The objectives of this study were: (1) to measure the effect of four instructional modes on the performance of tasks which embodied certain musical concepts, (2) to measure the performance differences between age levels, and (3) to measure and evaluate young children's concepts of louder, faster, higher, and shorter as elicited by the measuring instrument. Music educators such as Carlsen (1969) and Aronoff (1969) have emphasized the need for research on the development of musical concepts, while publishers and authors have produced music series and texts which espouse conceptual learning and teaching. Andrews and Deihl (1970) have recognized the need for research particularly with young children at the primary level while Zimmerman (1970) has shown the importance of developing instructional procedures for the attainment of musical concepts.

There are many definitions of the term "concept", however, general agreement exists concerning the importance of concepts in all intellectual functions. In this study a concept was defined as an internal structure whose existence is inferred whenever two or more discriminable objects or events have been grouped together and set apart from other objects or events on the basis of some common feature (Bourne, 1966).

Conceptual development involves a transition from perceptual dependence to inference. Three characteristics of conceptual development are: (1) the ability to identify the relevant feature within the perceptual field, i.e., to conserve or abstract the common element, (2) to require less redundant information as an aid in identifying the relevant feature, and (3) to discriminate between objects or events which are separated in either space or time, i.e., to remember relevant features even though the perceptual field changes. Concepts once attained tend to guide and organize subsequent perceptual activity. The role of language in concept development is probably most significant after primitive concepts have been developed. Language apparently facilitates the organization of primitive concepts into more complex structures.

The specific hypotheses tested in this study were:

> There is no significant difference between the observed mean on any task sequence and a mean which is attained merely by chance.
>
> There is no significant difference between the means attained by different age groups on any task sequence.
>
> There is no significant difference between the means attained by groups which are given different modes of instruction.

Two forms of the diagnostic instrument were developed and administered to independent samples of young children selected at random from public and parochial schools of Lincoln, Nebraska and the surrounding region. The Long form, which was given to 126 children, served as a prototype for the Short form which was given to 260 children. The children in the sample were selected from kindergarten, first, and second grade classes and were randomly placed into four nearly equal treatment groups. All groups were given the same orientation instructions; however, subsequent instructions varied according to the treatment group of which the child was a member. The modes of instruction for each group were (1) discovery, (2) verbal cue, (3) verbal response, and (4) motor response. Children were tested individually in four to eight sessions of about 20 minutes each. Both forms contained several tests each of which measured a particular concept. The Long form had eight different tests; the Short form only four. Each test consisted of 20 items or tasks with each task displaying positive and negative instances of the concept. The children were asked to identify the negative, i.e., "different", instance of the set. The tasks of the Long form consisted of four melodic patterns, three of which were positive instances and one which was negative. The size of the variation of the relevant and irrelevant features was determined by establishing a subjective scale of discriminability. The Short form tasks had only two instances, one positive, the other negative. Subjects were asked to identify the instance which was correct on the basis of a pre-test model task. Subjects who demonstrated consistent accuracy in their choices were assumed to have used a conceptual rule such as "the faster one."

Hypothesis One was intended to differentiate between conceptual behavior and chance performance. The results indicated significant differences in the level of performance depending on the type of concept embodied in the test. A concept of louder was well utilized by the majority of children at each grade level. The performance on tempo tasks was slightly lower but still significantly higher than chance expectation. On the duration tasks the children in the kindergarten and first grade groups did not show a consistent use of the concept of shorter; however, second grade performance indicated consistent conceptual behavior on these tasks. The lowest performance was achieved on pitch tasks where no group showed consistent use of a concept of higher. Nevertheless, within each group a number of individuals did perform at a level significantly higher than chance.

Hypothesis Two was that there is no significant difference between the means of different age groups. The results indicated that there were significant differences ($p < .01$) on each test between age levels. The greatest difference was between the kindergarten and the first grade performances with the second grade means being only somewhat higher than the first grade means on all tests except the pitch test where it was lower.

Hypothesis Three was that there is no significant difference in the means of the four treatment groups. The hypothesis was retained for each test except the tempo test. In some cases the mode of instruction was effective at one age level but not the others. In general, the verbal cue mode resulted in higher means than the non-verbal modes. The verbal cue mode resulted in significantly higher means in the case of the tempo tasks at each age level, on volume tasks at the kindergarten level, and on pitch and duration tasks at the second grade level. Only on the Long form pitch tasks was the mean of the verbal cue mode lower than the means of the non-verbal modes.

Conclusions

1. The level of performance on these tests is a function of the stimulus features of the tasks and the accessibility of the concept. A concept of louder appears to be well established and readily accessible. A concept of faster is also well established but slightly less so than that of louder. Duration and pitch concepts appear to be less accessible and more labile. The concept of higher is often confused with louder and children frequently apply the label "louder" to a higher pitch.

2. The level of performance on these tests is also a function of age. Whether the quantity of musical experience or a qualitative difference in children's thinking is responsible for the age differences cannot be determined with certainty. The discontinuity between the kindergarten and first grade means was interpreted as a possible manifestation of the shift from pre-operational to operational thought which, according th Piaget's theory, occurs about age seven.

3. There is usually no significant effect of the instructional mode when a concept is not readily accessible as in the case of the concept of higher. Neither is there a significant effect of the instructional mode when the concept is well developed as in the case of louder. On the other hand, a verbal mode of instruction, either

a verbal cue or a verbal response which is reinforced, is effective when the concept development is apparently at some intermediate stage as in the case of the concept of faster.

II. IMPLICATIONS FOR MUSIC EDUCATION

The results of this study imply certain emphases or strategies of teaching music at the primary level. These are:

A. Since children are likely to detect changes in volume more readily than other kinds of change, particularly pitch changes, it is necessary that volume be held constant, while pitch or other relevant features are allowed to vary. It is important, when asking children to compare sounds, to present stimulus materials which, at first, vary in only one relevant attribute. For example, if volume is an uncontrolled variable in tasks in which pitch levels are being compared, the two attributes may be confounded by the child.

B. Greater emphasis should be given to the development of pitch and durational concepts. Children are less likely to have attained these concepts than volume or tempo concepts. Even though the results of this study showed that pitch and durational concepts were less developed than the others, it should not be concluded that further improvement cannot result from carefully planned instruction.

C. In listening tasks where comparisons of different levels of a tonal feature, say, pitch, are being made, it is necessary that the contrast between the stimulus patterns be large enough to be readily perceived. At the initial stages of concept formation the sorting materials should be grossly different from each other, whereas at later stages the differences may become more subtle. With the development of stable concepts children may be expected to discriminate between ever smaller changes in tonal features. For example, pitch changes of an octave or more may be a suitable starting point in a problem designed to develop pitch concepts.

D. The verbal labels which teachers apply to tonal relationships should be readily related to the relevant attribute. It is important that the tonal referent be clearly specified. Inasmuch as tonal patterns are non-spatial and transitory, they should be represented as much as possible in spatial and more permanent form. The appropriate verbal label should become associated to the representation of the tonal pattern as well as with the tonal pattern itself.

Labels such as "high, up, down, soft, loud, faster, and short" should be applied to tonal features with accuracy and consistency. Obviously, a teacher should refrain from using such terms as "low" or "high" in reference to volume changes. There is also possible confusion over the term "fast" in reference to a durational and/or tempo change. Teachers should realize that the term they use to "explain" the changes in sound does not always have the same meaning to the child and may lead him to attend to the wrong feature.

E. The diagnosis of concept attainment is complicated by (1) the various manifestations of conceptual behavior, (2) the tendency to evaluate on the basis of the group response rather than individual behavior, and (3) the lack of objective, standardized measuring instruments. As a result, diagnosis is likely to be impressionistic and somewhat unreliable. It should be noted that:

1. The ability to sing or play tonal patterns is not in itself a sufficient indicator that concepts are being used by the child.

2. Verbal justification is useful as an indicator of concept attainment but the teacher must be careful to note whether the child's use of the term is consistent with its standard, public definition. Only after a child responds accurately and consistently over a series of tasks may it be concluded that the concept has been attained.

3. Non-verbal measures such as blank notation, hand or body movements are also useful indicators of concept attainment; however, their value in a group situation is handicapped by the difficulty of accurately determining correct and incorrect responses. Children readily imitate their teacher or their peers without really having the concept.

4. Individual assessment within the group is possible, but it is often impractical to record the observations in a systematic manner. Sorting games using tonal materials prepared for the Language Master, tape loops, or cassette may be useful both for learning and diagnosis of concept formation in situations where a listening center is available.

5. For effective teaching of concepts the teacher needs to use some means of evaluating individual musical growth. For the present these procedures will have to be "home-made" since no published instruments have been produced.

III. RECOMMENDATIONS FOR FURTHER RESEARCH

A. There is a need to determine the effect of task variables on conceptual behavior. Among these variables are:

1. Varying amounts of redundant information in the stimulus pattern. For example, a comparison may be made of the effect of two kinds of tasks, one which contains redundant information, the other no redundant information. In the first kind of task a

relevant attribute, say, pitch, may be combined with timbre or other attribute. Either of these attributes may be used to solve a sorting problem. In the second kind of task there would be only one relevant attribute.

2. Varying amounts of irrelevant, competing information in the stimulus pattern. For example, tasks may involve pitch as a relevant attribute but also have other irrelevant aspects such as tempo, duration, and so on. These tasks differ from those described above since the irrelevant attribute competes for attention and does not lead to a correct solution of the problem. It would also be necessary to have tasks which had no irrelevant information in them as a basis of comparison.

3. Varying amounts of change of the relevant attribute between negative and positive instances. For example, instead of varying the pitch between the positive and negative instances by the interval of a third, various intervals may be used. Other features also may be changed by varying amounts.

B. A continued study of the relationship between verbal and non-verbal instructional modes is encouraged.

C. There is a need for technical equipment for presenting tonal materials which allows for flexibility and manipulability by the subject. In sorting tasks it is highly desirable to allow the subject to re-order the stimulus patterns or to vary the speed of presentation. A device having the flexibility of the Language Master with improved fidelity would be useful for this purpose.

D. There is a need for continued investigation of the development of musical concepts across age levels. Of particular interest is the formative stage in the pre-school period. Such research may begin with naturalistic observation and analysis of behavior which is later followed by controlled experimentation.

E. There is a need to develop intensive training studies which are designed to develop or strengthen specific concepts. These studies should begin with a pre-test of conceptual performance. A possible strategy is that of the oddity problem which was described in Chapter Two. This strategy allows the subject to repeat a task until a high level of accuracy is achieved before a new task is taken up. An earlier study by the investigator indicated that such a strategy may be successful in the development of musical concepts, but a more careful application of this strategy is needed.

REFERENCES

Andrews, Frances M., and Deihl, N. C. "Development of a Technique for Identifying Elementary School Children's Musical Concepts." Council for Research in Music Education, 1968, Bulletin No. 13, 1-7.

Andrews, Frances M., and Deihl, N. C. "Development of a Technique for Identifying Elementary School Children's Musical Concepts." Journal of Research in Music Education, 1970, 18, No. 3, 214-222.

Aronoff, Frances. Music and Young Children. New York: Holt, Rinehart & Winston, 1969.

Ausubel, D. P. "In Defense of Verbal Learning." In R. C. Anderson and D. P. Ausubel (Eds.), Readings in the Psychology of Cognition. New York: Holt, Rinehart & Winston, 1965, Pp. 87-102.

Baldwin, A. L. Theories of Child Development. New York: Wiley, 1969.

Benn, Oleta A. "A Message for Teachers." In N. B. Henry (Ed.), Basic Concepts in Music Education. Fifty-seventh Yearbook of the National Society for the Study of Education, Part I. Chicago: University of Chicago Press, 1958. Pp. 339-355.

Boring, E. G. A History of Experimental Psychology. (2nd ed.) New York: Appleton-Century-Crofts, 1950.

Bourne, L. E. Human Conceptual Behavior. Boston: Allyn & Bacon, 1966.

Braine, M. D. "Ontogeny of Certain Logical Operations: Piaget's Formulation Examined by Nonverbal Methods." In I. E. Sigel and F. H. Hooper (Eds.), Logical Thinking in Children. New York: Holt, Rinehart & Winston, 1968. Pp. 164-206.

Broadbent, D. E. Perception and Communication. London: Perganmon Press, 1958.

Bruner, J. S., Goodnow, J. J., and Austin, G. A. A Study of Thinking. New York: Wiley, 1956.

Bruner, J. S. "On Perceptual Readiness." Psychological Rev., 1957, 64, 123-152.

Bruner, J. S. "On Going Beyond the Information Given." In R. J. Harper et al. (Eds.), The Cognitive Processes: Readings. Englewood Cliffs, New Jersey: Prentice-Hall, 1964. Pp. 293-310.

Bruner, J. S., Olver, Rose R., and Greenfield, Patricia M. (with others). Studies in Cognitive Growth. New York: Wiley, 1966.

Brunswik, E. Perception and the Representative Design of Psychological Experiments. (2nd ed.), Berkeley, California: University of California Press, 1956.

Brunswik, E. "Reasoning as a Universal Behavior Model and a Functional Differentiation between 'Perception' and 'Thinking'." In R. Hammond (Ed.), The Psychology of Egon Brunswik. New York: Holt, Rinehart & Winston, 1966. Pp. 487-493.

Carabo-Cone, Madeline. A Sensory-Motor Approach to Music Learning. Book 1. New York: MCA Music, 1969.

Carlsen, J. C. "Some Problems in Musical Learning." Journal of Research in Music Education, 1969, 17, No. 1, 7-12.

Carroll, J. B. Language and Thought. Englewood Cliffs, New Jersey: Prentice-Hall, 1964. (a)

Carroll, J. B. "Words, Meanings, and Concepts." Harvard Educational Review, 1964, 34, 178-202. (b)

Cassirer, E. The Logic of the Humanities. New Haven and London: Yale University Press, 1960.

Cassirer, E. "On the Theory of the Formation of Concepts." In L. Uhr (Ed.), Pattern Recognition. New York: Wiley, 1966. Pp. 20-34.

Charlesworth, W. R. "Cognition in Infancy." In Stella Chess and T. Alexander (Eds.), Annual Progress in Child Psychiatry and Child Development. New York: Brunner/Mazel, 1969. Pp. 1-23.

Cheyette, I. and H. Teaching Music Creatively in the Elementary School. New York: McGraw, 1969.

Communication Research Machines, Inc. Psychology Today. Del Mar, California: CRM Books, 1970.

Communication Research Machines, Inc. Developmental Psychology Today. Del Mar, California: CRM Books, 1971.

Dale, E. "Education as Conceptualizing." The News Letter, 1971, 36, No. 6, 1-4.

Farnsworth, P. R. The Social Psychology of Music. (2nd ed.) Ames, Iowa: Iowa State University Press, 1969.

Flavell, J. H. The Developmental Psychology of Jean Piaget. Princeton, New Jersey: Van Nostrand, 1963.

Flavell, J. H., and Wohlwill, J. F. "Formal and Functional Aspects of Cognitive Development." In D. Elkind and J. H. Flavell (Eds.), Studies in Cognitive Development: Essays in Honor of Jean Piaget. New York: Oxford University Press, 1969. Pp. 67-120.

Flavell, J. H. "Concept Development." In P. H. Mussen (Ed.), Carmichael's Manual of Child Psychology, Vol. 1, (3rd ed.) New York: Wiley, 1970, Pp. 983-1060.

Furth, H. G. Thinking Without Language: Psychological Implications of Deafness. New York: Free Press, 1966.

Furth, H. G. Piaget and Knowledge: Theoretical Foundations. Englewood Cliffs, New Jersey: Prentice-Hall, 1969.

Gagne', R. M. The Conditions of Learning. New York: Holt, Rinehart & Winston, 1965.

Gange', R. M. "The Learning of Principles." In H. Klausmeier (Ed.), Analyses of Concept Learning. New York: Academic Press, 1966. Pp. 81-87.

Garrett, H. E. Statistics in Psychology and Education. (5th ed.) New York: Longmans, Green and Co., 1958.

Gary, C. L. (Ed.) The Study of Music in the Elementary School - A Conceptual Approach. Washington, D. C.: Music Educators National Conference, 1967.

Guilford, J. P. Fundamental Statistics in Psychology and Education. (4th ed.) New York: McGraw-Hill, 1956.

Guilford, J. P., Dunham, J. L., and Hoepfner, R. "Abilities Pertaining to Classes and the Learning of Concepts." Reports from the Psychological Laboratory, Los Angeles: University of Southern California, 1966, No. 39.

Guilford, J. P. The Nature of Human Intelligence. New York: McGraw-Hill, 1967.

Harvey, O. J., Hunt, D. E., and Schroder, H. M. Conceptual Systems and personality organization. New York: Wiley, 1961.

Harlow, H. F. "Learning Set and Error Factor Theory." In S. Koch (Ed.), Psychology, A Study of a Science. Vol. I. New York: McGraw-Hill, 1959.

Harlow, H. F. "The Formation of Learning Sets." Psychological Review, 1949, 56, 51-65.

Hickman, A. T. "Some Preverbal Concepts in Music." Journal of Research in Music Education, 1969, 17, No. 1, 90-93.

Hunt, E. B. Concept Learning: An Information Processing Problem. New York: Wiley, 1962.

Kagan, J. and Kogan, N. "Individuality and Cognitive Performance." In P. H. Mussen (Ed.), Carmichael's Manual of Child Psychology. Vol. 1. (3rd ed.) New York: Wiley, 1970. Pp. 1273-1378.

Kendler, T. S. "Concept Formation." Annual Review of Psychology, 1961, 13, 447-472.

Kessen, W. "Research Design in the Study of Developmental Problems." In P. H. Mussen (Ed.), Handbook of Research Methods in Child Development. New York: Wiley, 1960. Pp. 36-70.

Kidd, Aline H. and Rivoire, J. L. (Eds.) Perceptual Development in Children. New York: International Universities Press, 1962.

Klausmeier, H. J., et al. Concept Learning and Problem Solving: A Bibliography, 1950-1964. Technical Report No. 1, 1965, University of Wisconsin, Research and Development Center for Learning and Re-education.

Koffka, Kurt. The Growth of the Mind. New York: Harcourt-Brace, 1924.

Kohlberg, L. "Early Education: A Cognitive-Developmental View." In Stella Chess and T. Alexander (Eds.), Annual Progress in Child Psychiatry and Child Development. New York: Brunner/Mazel, 1969, Pp. 72-124.

Kowall, Bonnie C. (Ed.) Perspectives in Music Education - Source Book III. Washington, D. C.: Music Educators National Conference, 1966.

Larsen, R. L., and Boody, C. G. "Some Implications for Music Education in the Work of Jean Piaget." Journal of Research in Music Education, 1971, 19, No. 1, 35-50.

Laurendeau, M., and Pinard, A. Causal Thinking in the Child. New York: International Universities Press, 1962.

Lehman, P. R. Tests and Measurements in Music. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

Leonard, C. and House R. W. Foundations and Principles of Music Education. New York: McGraw-Hill, 1959.

Leonard, C., Krone, Beatrice, Wolfe, I., and Fullerton, Margaret. Discovering Music Together. Book 1, Teacher's Edition. Chicago: Follett, 1967.

Lundin, R. W. An Objective Psychology of Music. New York: Ronald Press, 1953.

Maccoby, Eleanor E., and Konrad, K. W. "The Effect of Preparatory Set on Selective Listening: Developmental Trends." Monographs of the Society for Research in Child Development, 1967, 32, No. 4, (Whole No. 112).

Madsen, C. K., and Madsen, C. H. Experimental Research in Music. Englewood Cliffs, New Jersey: Prentice-Hall, 1970.

Marsh, Mary V. Explore and Discover Music. New York: MacMillan, 1970.

McKee, J. P., and Riley, D. A. "Auditory Transposition in Six-Year Old Children." Child Development, 1962, 33, No. 2, 469-476.

Melton, A. W. Categories of Human Learning. New York: Academic Press, 1964.

Meyer, L. B. Music, the Arts, and Ideas. Chicago: University of Chicago Press, 1967.

Mursell, J. L. The Psychology of Music. New York: W. W. Norton & Co., 1937.

Neisser, U. Cognitive Psychology. New York: Appleton-Century-Crofts, 1967.

Nye, R. E., and Nye, Vernice T. Music in the Elementary School. (3rd ed.) Englewood Cliffs, New Jersey: Prentice-Hall, 1970.

Pezold, R. G. "Auditory Perception by Children." Journal of Research in Music Education, 1969, 17, No. 1, 82-87.

Pflederer, Marilyn. "The Response of Children to Musical Tasks Embodying Piaget's Principle of Conservation." Journal of Research in Music Education, 1964, 12, No. 4, 251-265.

Pflederer, Marilyn. "How Children Conceptually Organize Musical Sounds." Council for Research in Music Education. 1966, Bulletin No. 7, 1-12.

Pflederer, Marilyn. "Conservation Laws Applied to the Development of Musical Intelligence." Journal of Research in Music Education, 1967, 15, No. 3, 215-223.

Pflederer, Marilyn, and Sechrest, L. "Conservation Type Responses of Children to Musical Stimuli." Council for Research in Music Education, 1968, Bulletin No. 13, 19-36.

Phenix, P. H. Philosophy of Education. New York: Holt, Rinehart & Winston, 1958.

Piaget, J. The Origins of Intelligence in Children. New York: W. W. Norton & Co., 1952.

Piaget, J., and Inhelder, B. The Psychology of the Child. New York: Basic Books, 1969.

Piaget, J. The Mechanisms of Perception. New York: Basic Books, 1969.

Piaget, J. "Piaget's Theory." In P. H. Mussen (Ed.), Carmichael's Manual of Child Psychology, Vol. 1. (3rd ed.) New York: Wiley, 1970. Pp. 703-732.

Piaget, J. Structuralism. New York: 1970.

Reimer, B. A Philosophy of Music Education. Englewood Cliffs, New Jersey: Prentice-Hall, 1970.

Riley, D. A., and McKee, J. P. "Pitch and Loudness Transposition In Children and Adults." Child Development, 1963, 34, No. 2, 471-482.

Sachs, K. The Wellsprings of Music. New York: McGraw Hill, 1965.

Shuter, Rosamund. The Psychology of Musical Ability. London: Methuen & Co., 1968.

Shuter, Rosamund. "Some Problems in the Psychology of Musical Ability." Journal of Research in Music Education, 1969, 17, No. 1, 90-93.

Sigel, I. E. "The Attainment of Concepts." In M. L. Hoffman and Lois W. Hoffman (Eds.), Review of Child Development Research. Vol. 1. New York: Russel Sage Foundation, 1964. Pp. 209-248.

Sigel, I. E., and Hooper, F. H. (Eds.) Logical Thinking in Children. New York: Holt, Rinehart, and Winston, 1968.

Stevens, S. S., and Davis, H. Hearing: Its Psychology and Physiology. New York: Wiley, 1938.

Sur, W. R., Tolbert, Mary R., Fischer, W., and McCall, Adeline. This is Music. Book 2, Teacher's Manual. Boston: Allyn and Bacon, 1967.

Underwood, B. J. Experimental Psychology (2nd ed.) New York: Appleton-Century-Crofts, 1966.

Vernon, M. D. Perception Through Experience. New York: Barnes & Noble, 1970.

Watters, Lorrain, et al. The Magic of Music. Kindergarten Book. Boston: Ginn & Co., 1965.

White, S. B. "Learning." In H. W. Stevenson (Ed.), Child Psychology. Sixty-second Yearbook of the National Society for the Study of Education, Part I. Chicago: University of Chicago Press, 1963. Pp. 196-235.

White, S. "Evidence for a Hierarchical Arrangement of Learning Processes." In L. P. Lipsitt and C. C. Spiker (Eds.), Advances in Child Development and Behavior. Vol. 2. New York: Academic Press, 1965. Pp. 187-216.

White, S. "The Learning Theory Approach." In P. H. Mussen (Ed.) Carmichael's Manual of Child Psychology. Vol. 1. (3rd ed.) New York: Wiley, 1970, Pp. 657-701.

Wohlwill, J. F. "From Perception to Inference: A Dimension of Cognitive Development." In I. E. Sigel and F. H. Hooper (Eds.), Logical Thinking in Children. New York: Holt, Rinehart & Winston, 1968. Pp. 472-494.

Woodruff, A. D. Basic Concepts of Teaching. San Francisco: Chandler Publishing Co., 1961.

Woodruff, A. D. "How Music Concepts Are Developed." Music Educators Journal, 1970, 56, No. 6, 51-54.

Worf, B. L. Language, Thought, and Reality. Cambridge and New York: M.I.T. - Wiley, 1956.

Zimmerman, Marilyn Pflederer. "Brief Focused Instruction and Musical Concepts." Journal of Research in Music Education, 1970, 18, No. 1, 25-36.

Addendum

Archer, E. J. "On Verbalization and Concepts." In A. W. Melton (Ed.), Categories of Human Learning. New York: Academic Press, 1964.

Pflederer (Zimmerman), Marilyn. "Brief Focused Instruction and Musical Concepts." Journal of Research in Music Education, 1970, 18, No. 1, 25-36.

Woodruff, A. D. "Concept Teaching in Music." In Bonnie C. Kowall (Ed.), Perspectives in Music Education - Source Book III. Washington, D. C.: Music Educators National Conference, 1966.

List of Appendixes

Appendix I: Melodic Patterns Used in Long and Short Forms 97

Appendix II: Long Form Results: Percentage of Subjects Reaching the Criterion on Each Test of the Long Form 98

Appendix III: Summary of Analyses of Variance Results From Long Form 99

Appendix IV: Long Form Means by Sequence, Grade, and Mode Five Tasks 100

Appendix V: Long Form Means by Sequence, Grade, and Mode Ten Tasks 101

Appendix VI: Summary of Short Form Means and Standard Deviations by Mode, Sequence, and Grade 102

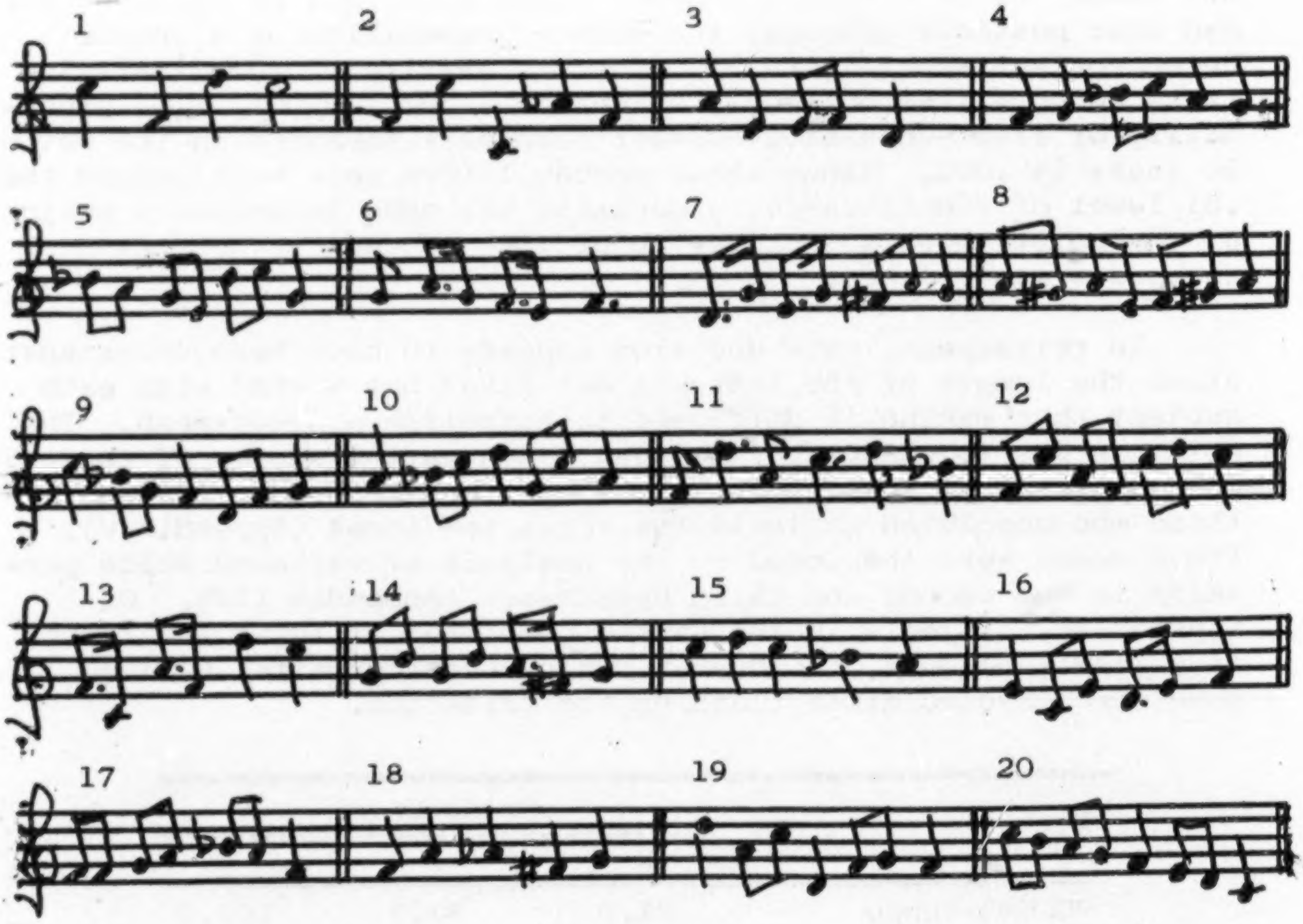
Appendix VII: Short Form Item Analysis 103

Appendix VIII: Sample Letter to Parents Whose Child Is Selected for the Music Research Project 104

Appendix IX: Instructions for Administering Tests 105

APPENDIX I

MELODIC PATTERNS USED IN LONG AND SHORT FORMS


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

APPENDIX II

LONG FORM RESULTS: PERCENTAGE OF SUBJECTS REACHING THE CRITERION ON EACH TEST OF THE LONG FORM

Note: A summary of the data from the Long form tests is found in Appendix II, III, IV, and V. Appendix II shows the percentage of subjects reaching the criterion on each test, and pertains to the first hypothesis, i.e., there are no significant differences between the performance levels on the various tests. The statistical procedure used with the Long form to test this hypothesis was based on the binomial model. Since each task of the Long form had four possible choices, the chance probability of a choice being correct was 1/4 or .25. The probability of the first five tasks being solved correctly by chance alone is .001. The probability of five successive correct responses anywhere in the set of 20 tasks is .004. Since these probabilities were well beyond the .01 level of significance, a decision was made to excuse a subject at the point where the criterion of five successive correct responses was reached.

In retrospect, this decision appears to have been unfortunate since the length of the test was not fixed but varied with each subject thus making it difficult to establish a group mean. The solution was to calculate the mean based on the scores of those Ss who completed at least the first five items (Appendix IV) and those who completed at least the first ten items (Appendix V). These means were then used in the analysis of variance which pertains to the second and third hypotheses (Appendix III). Of course, all subjects in the sample completed at least the first five items; however, the number diminished thereafter because some were excused after reaching the criterion.

Test	Grade Level		
	K	1	2
VOLUME-tempo	81.0	81.1	100.0
VOLUME-pitch	95.2	100.0	97.6
TEMPO-volume	9.5	38.1	40.5
TEMPO-duration	26.2	59.5	47.6
PITCH-duration	4.8	50.0	31.0
PITCH-volume	4.8	52.4	28.6
DURATION-pitch	16.7	45.2	61.9
DURATION-volume	47.6	90.5	83.3

Upper case = Relevant attribute
Lower case = Irrelevant attribute

APPENDIX III

SUMMARY OF ANALYSES OF VARIANCE RESULTS FROM LONG FORM

Note: Whenever a subject reached the criterion of five successive correct responses, he was excused from further work on the test. As a result, the number of subjects who completed any item beyond the fifth did not remain constant. An analysis of variance was made by considering the fifth and the tenth item as end-point of the test. The scores of all subjects who worked up to and including the fifth and tenth items were used in calculating the variance. In the summaries below, the length of the test on which the analysis was made is shown in parentheses. In certain tests so many Ss were excused before the tenth item was reached that an analysis of variance was very unreliable since so few Ss were involved. No report is given in these cases.

Test		Effect	DF_1	DF_2	F.Ratio	Significance
VOLUME-tempo	(5)	Age	2	114	2.27	ns
		Mode	3	114	4.00	$p < .01$
VOLUME-pitch	(5)	Age	2	114	0.99	ns
		Mode	3	114	0.74	ns
TEMPO-volume	(5)	Age	2	114	13.43	$p < .01$
		Mode	3	114	2.18	ns
TEMPO-volume	(10)	Age	2	103	9.53	$p < .01$
		Mode	3	103	2.68	ns
TEMPO-duration	(5)	Age	2	114	4.68	$p < .01$
		Mode	3	114	2.91	$p < .05$
TEMPO-duration	(10)	Age	2	69	5.66	$p < .01$
		Mode	3	69	1.43	ns
PITCH-duration	(5)	Age	2	114	12.05	$p < .01$
		Mode	3	114	0.78	ns
PITCH-duration	(10)	Age	2	92	8.22	$p < .01$
		Mode	3	92	1.88	ns
PITCH-volume	(5)	Age	2	114	2.88	ns
		Mode	3	114	0.35	ns
PITCH-volume	(10)	Age	2	96	1.73	ns
		Mode	3	96	1.18	ns
DURATION-pitch	(5)	Age	2	114	6.31	$p < .01$
		Mode	3	114	0.80	ns
DURATION-pitch	(10)	Age	2	75	6.25	$p < .01$
		Mode	3	75	0.47	ns
DURATION-volume	(5)	Age	2	114	13.08	$p < .01$
		Mode	3	114	0.51	ns

APPENDIX IV

LONG FORM MEANS BY SEQUENCE, GRADE, AND MODE

Number of subjects completing five tasks is shown in parentheses. Upper case = relevant feature, lower case = irrelevant feature.

Sequence	Grade	Mode			
		1	2	3	4
VOLUME tempo	K	3.44(9)	4.23(13)	3.00(10)	3.40(10)
	1	3.60(10)	4.45(11)	3.80(10)	3.45(11)
	2	3.67(12)	4.80(10)	4.00(10)	4.20(10)
TEMPO volume	K	1.11(9)	1.38(13)	1.70(10)	0.90(10)
	1	2.00(10)	2.82(11)	1.60(10)	2.09(11)
	2	2.42(12)	3.50(10)	2.40(10)	2.30(10)
PITCH duration	K	1.33(9)	1.85(13)	1.30(10)	1.20(10)
	1	2.80(10)	2.55(11)	2.00(10)	2.55(11)
	2	3.08(12)	2.60(10)	2.60(10)	3.00(10)
DURATION pitch	K	2.22(9)	2.46(13)	2.40(10)	2.50(10)
	1	3.60(10)	2.27(11)	3.20(10)	3.36(11)
	2	3.00(12)	3.40(10)	3.40(10)	3.50(10)
VOLUME pitch	K	4.44(9)	4.15(13)	4.30(10)	3.90(10)
	1	4.30(10)	4.73(11)	4.40(10)	4.45(11)
	2	4.42(12)	5.00(10)	3.90(10)	4.60(10)
PITCH volume	K	1.67(9)	1.31(13)	1.90(10)	2.10(10)
	1	3.00(10)	2.36(11)	2.00(10)	2.18(11)
	2	2.08(12)	2.40(10)	2.20(10)	2.40(10)
TEMPO duration	K	2.44(9)	2.62(13)	3.00(10)	1.90(10)
	1	3.20(10)	4.18(11)	2.80(10)	3.09(11)
	2	2.58(12)	4.00(10)	2.70(10)	3.20(10)
DURATION tempo	K	2.78(9)	3.15(13)	2.50(10)	2.20(10)
	1	3.80(10)	3.55(11)	3.70(10)	3.91(11)
	2	4.33(12)	3.50(10)	4.00(10)	4.90(10)

APPENDIX V

LONG FORM MEANS BY SEQUENCE, GRADE, AND MODE

Number of subjects completing ten tasks is shown in parentheses.
Upper case = relevant feature, lower case = irrelevant feature.

Sequence	Grade	Mode			
		1	2	3	4
VOLUME tempo	K	6.25(4)	5.00(2)	5.57(7)	4.75(4)
	1	7.00(1)	4.50(2)	4.00(2)	6.00(5)
	2	4.60(5)	0.00(0)	6.00(1)	6.50(4)
TEMPO volume	K	2.33(9)	2.77(13)	3.00(10)	2.40(10)
	1	3.63(8)	5.55(11)	2.89(9)	3.90(10)
	2	4.33(12)	5.29(7)	3.25(8)	4.13(8)
PITCH duration	K	3.78(9)	3.46(13)	3.20(10)	2.70(10)
	1	4.63(8)	4.14(7)	4.00(6)	5.00(6)
	2	5.55(11)	3.63(8)	4.50(8)	5.50(8)
DURATION pitch	K	3.56(9)	3.50(10)	3.33(9)	4.22(9)
	1	4.20(5)	4.33(9)	5.14(7)	5.33(6)
	2	5.57(7)	5.75(4)	5.17(6)	5.50(6)
VOLUME pitch	K	4.00(1)	6.67(3)	6.00(1)	3.00(1)
	1	7.00(2)	7.00(1)	3.00(1)	0.00(0)
	2	4.00(1)	0.00(0)	0.00(1)	0.00(0)
PITCH volume	K	3.22(9)	2.54(13)	3.30(10)	3.00(10)
	1	3.00(6)	2.60(5)	4.00(10)	4.67(9)
	2	3.45(11)	3.33(9)	2.88(8)	3.50(8)
TEMPO duration	K	5.00(7)	4.64(11)	5.29(7)	3.25(8)
	1	4.29(7)	5.00(3)	4.20(5)	3.75(4)
	2	5.30(10)	6.83(6)	4.29(7)	6.33(6)
DURATION tempo	K	4.71(7)	5.25(8)	3.88(8)	3.63(8)
	1	6.25(4)	6.14(7)	5.00(3)	4.67(3)
	2	5.50(4)	5.50(4)	5.00(4)	7.00(1)

APPENDIX VI

SUMMARY OF SHORT FORM MEANS AND STANDARD DEVIATIONS BY MODE, SEQUENCE, AND GRADE

		Mode							
Sequence	Grade	1		2		3		4	
		Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Volume	K	15.33	(4.35)	17.10	(2.90)	15.39	(4.50)	13.73	(4.06)
	1	17.90	(3.79)	18.29	(2.51)	17.95	(3.02)	18.61	(2.23)
	2	18.14	(2.95)	18.78	(1.51)	18.45	(2.13)	18.22	(3.00)
Tempo	K	12.33	(3.69)	14.75	(4.06)	13.30	(4.05)	12.27	(3.67)
	1	14.81	(4.33)	16.95	(3.19)	17.85	(1.90)	16.96	(3.13)
	2	16.19	(3.70)	18.30	(1.89)	17.00	(3.35)	15.87	(2.74)
Pitch	K	10.38	(2.85)	11.10	(4.15)	11.09	(2.47)	10.77	(3.48)
	1	12.29	(3.70)	11.38	(4.49)	13.55	(3.50)	13.39	(4.32)
	2	9.90	(2.45)	13.13	(4.60)	13.36	(4.32)	10.61	(3.49)
Duration	K	11.76	(3.28)	12.55	(3.90)	12.04	(4.08)	12.00	(3.06)
	1	13.81	(3.86)	14.10	(3.66)	13.25	(4.85)	12.43	(4.58)
	2	14.10	(3.67)	15.74	(3.08)	13.32	(4.24)	13.61	(4.24)

APPENDIX VII

SHORT FORM ITEM ANALYSIS

Item	Frequency of Correct Responses											
	Volume Test			Tempo Test			Pitch Test			Duration Test		
	K	1	2	K	1	2	K	1	2	K	1	2
1	64	76	80	64	78	79	52	58	53	57	58	63
2	59	75	79	66	74	73	41	50	50	60	67	69
3	67	82	83	51	71	73	32	38	47	45	54	62
4	68	80	80	60	71	83	47	57	58	57	59	68
5	63	75	88	58	70	78	52	50	51	57	52	62
6	55	78	79	55	68	72	53	58	58	45	49	58
7	65	77	81	45	67	63	56	59	59	50	48	56
8	71	73	84	55	71	74	43	57	64	52	53	56
9	62	76	79	44	54	55	50	54	52	55	62	60
10	67	78	80	58	74	75	50	56	51	60	56	62
11	69	78	80	57	78	80	43	49	47	45	59	60
12	69	76	84	61	80	82	50	56	49	63	59	71
13	70	80	82	58	66	71	51	58	48	41	52	54
14	66	74	81	56	65	76	45	52	51	46	62	65
15	72	78	80	59	72	77	52	58	56	55	57	72
16	72	76	86	55	61	77	47	57	55	48	53	62
17	67	79	85	52	79	80	48	48	43	50	57	67
18	68	79	85	65	76	84	40	52	48	57	61	65
19	62	78	78	58	71	78	35	57	56	51	51	62
20	64	79	84	53	81	69	45	52	54	45	61	70

Total N's by grade: K=86, 1=85, 2=89

APPENDIX VIII

SAMPLE LETTER TO PARENTS WHOSE CHILD IS SELECTED FOR THE MUSIC RESEARCH PROJECT

Dear Parents:

Our school has been asked to participate in a research study on the development of musical concepts in young children. We believe that the study has merit and may lead to a better understanding of how children learn. The project is supported by the U.S. Office of Education and is being directed by Donald Taebel, Professor of Music at Concordia Teachers College.

Several children from the kindergarten and first grade are being considered for this study. We would like your child to also be included in the project. Your child's teacher already knows of his selection and has given her approval.

In brief, each child in the project will be interviewed twice a week for four or five weeks. During the interview he will be asked to listen to musical examples and discover various tonal changes in these examples. By observing a child's response to these musical "puzzles", we hope to learn not only what his concepts of music are, but also how he learns new concepts. In other words, we are not only interested in what your child already knows, but also how he learns something new. A report of your child's achievement in this study will be provided if you request it.

If you approve of your child being in the project please return the attached permission form as soon as possible. Should you have any questions concerning the project, please call Mr. Taebel at 643-3381 or 643-3651, ext. 256.

Eloise Busche
Elementary Principal

(Please return to the Classroom Teacher)

I approve of my child taking part in the research project on the development of musical concepts.

____________________________ __________
Parent's Signature Date

Child's Name______________________________ Birthdate_______ _____
Date Year

APPENDIX IX

INSTRUCTIONS FOR ADMINISTERING TESTS

Note: See page 51 for orientation instructions.

CONDUCTING THE INDIVIDUAL INTERVIEW. After this point has been reached say, "Now that you know how to play the listening game, I would like to have each of you try some other games by yourself. You may all go back to your room now and I will call for you one at a time." Before the children leave explain that after each child is done he will go back to the room and tell the next one to come. Arrange the order now. Go back to the room with the children and ask one of them to come back to the test area with you.

The individual instructions will vary depending on the response mode to which a child has been assigned. There are four response modes. An explanation of the response modes is given on the accompanying colored sheets. The response mode is indicated on each child's white data form. Following the data forms there should be four colored sheets. The color of these forms indicates the mode of response, i.e., green-mode 1, yellow-mode 2, blue-mode 3, and pink-mode 4.

There are four concepts being measured; they are sequence A-loudness, sequence B-faster, sequence C-higher, and sequence D-shorter. Each child will take four tests on these concepts. The tests are given in a mixed order to the children. One child may take the tests in ABCD order, another may take them in DCBA order, and so on. The order in which the tests should be given is indicated by the capital letters on the white sheet of each child. The letters SF stand for "short form".

First, you must note the response mode of the child, i.e., 1,2, 3, or 4 as indicated on the white sheet. Then you must note which test sequence the child is to take. Having determined these two things, you are now ready to proceed with the individual test instructions.

Let us assume that the child is assigned to response mode 1 and that he is to start with test C, which is a test of the concept "higher". Say, "Do you remember the listening game we played last time? You had to find out what kind of tune I liked best. Let me play a tune for you again. (Play first phrase of 'Mary Had a Little Lamb'.) Now I'm going to change the tune a little, listen. (Play the phrase again, but higher.) Did you hear the change? I like the second way better. Let's try the game without my helping you and you listen to the kind of change I like." (Play phrases again, the same order.) If the child gets this right, repeat the phrases with the higher one being either first or second. Do this until the child has three correct responses. Next say, "Now listen to the tunes on the tape recorder. Listen carefully for the change in the tune which I like. It could be either the first or second tune."

Play the taped sequence on the cassette. When the first example is being played hold up one finger of left hand, hold up two fingers of your right hand for second example. After the second example, stop the tape and ask the child which one he thinks you like the best. If he is correct, say, "Yes, that's the right one, etc." and give him a chip. If he gives the wrong answer, say, "No, it was the other tune." Mark his choice number 1 or 2 and indicate in the box below whether it was correct with a plus or incorrect with a 0.

RESPONSE MODE 2: (Example of instructions for duration test) Same introductory remarks as in mode 1. After saying, "Do you hear the change? I like that way better" add "I like the tune when the notes are shorter." Then continue with instructions as before to this point: "Now listen to the tunes on the tape recorder. I like the tune where the notes are shorter. Listen for that kind of change. It might be either the first or second tune."

For other sequences than D, in which the concept is shorter, say "louder" for A, "faster" for B, and "higher" for C. If a child asks what those words mean give him a brief, simple definition and ask him to listen to your example. Do not show the meaning of these words with any movement or drawing. Indicate such questions on the data form. You should remind the child after each game that he is listening for the tune where the notes are shorter, etc.

RESPONSE MODE 3: Give the same general instructions as in Mode 1. After the first game, ask the child, "Do you know what kind of change I like?" The possible responses are these:
a. He may choose the correct tune, but give you a wrong or ambiguous answer. Write down his wrong answer for each game.
b. He may choose the correct tune, and give you a correct answer. If he says louder, faster, higher, or shorter show this answer with a plus in the verbal response section. If he gives you a synonym such as jumpy, bumpy, space between the sounds, etc., write this down.
c. If he gets the wrong choice, ask him if he remembers what kind of change you like. Remind him to listen for it. Ask him to verbalize only if he gets the right answer.

Note: The "correctness" of the response is based on his initial choice, NOT his verbal answer. Therefore give him a chip even if he gives you a wrong verbal answer. Mark the data sheet in the boxes + or 0 for the initial choice, not his verbal answer. Use the section provided for that answer.

RESPONSE MODE 4: Motor response. Have child tap or hum along with you as you sing the examples. Have him tap for sequences A, B, and D. Have him hum or move his hand to reflect pitch changes for sequence C. If the child does not respond, you should do the tapping, humming, or moving your hand.

Indicate on the data sheet which motor activity the child prefers.